

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 22:32:51 ; Search time 2816 Seconds
(without alignments)
2966.089 Million cell updates/sec

Title: US-09-925-637-64

Perfect score: 287

Sequence: 1 MTNLPKMKLIDEVNNELSA.....ELTQIDQFNKHLLEIVDL 287

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 2054640 seqs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174850_24112/app_query.fasta_1.455
-DB=GenEmbl -OFMT=fastap -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIS-bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637.ecgn_1_1_2496_runat_23052003_174850_24112 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg_hum:*
31: em.htg_inv:*
32: em.htg_other:*
33: em.htg_mus:*
34: em.htg_pln:*
35: em.htg_rtd:*
36: em.htg_mam:*
37: em.htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	197	68.6	333750	1	AP004827	AP004827 Staphyloc
C 2	132	46.0	301550	1	AP003134	AP003134 Staphyloc
C 3	132	46.0	346900	1	AP003362	AP003362 Staphyloc
4	102	35.5	882	6	AR106466	AR106466 Sequence
5	37	12.9	783	6	AX141501	AX141501 Sequence
6	37	12.9	3444	1	AF270104	AF270104 Staphyloc
7	37	12.9	3444	6	AX145422	AX145422 Sequence
C 8	37	12.9	3454	1	AF269958	AF269958 Staphyloc
C 9	37	12.9	3454	6	AX145276	AX145276 Sequence
C 10	37	12.9	4045	1	AF269889	AF269889 Staphyloc
C 11	37	12.9	4045	6	AX145207	AX145207 Sequence
12	33	11.5	486	6	AR097529	AR097529 Sequence
13	33	11.5	486	6	AR142282	AR142282 Sequence
14	33	11.5	486	6	EL5319	EL5319 Micrococcus
15	33	11.5	1080	1	AB003187	AB003187 Micrococc
16	27	9.4	2268	1	LM0012349	LM0012349 Listeria
17	27	9.4	5635	6	AX146832	AX146832 Sequence
18	27	9.4	250050	1	AL591978	AL591978 Listeria
19	19	6.6	957	1	AB016093	AB016093 Synechoco
C 20	19	6.6	195269	6	AX417035	AX417035 Sequence
C 21	19	6.6	300600	1	AP005369	AP005369 Thermosyn
22	19	6.6	333050	1	AL596168	AL596168 Listeria
23	19	6.6	349980	6	AX417044	AX417044 Sequence
24	18	6.3	10029	1	AE006320	AE006320 Lactococc
25	18	6.3	10029	1	AE013090	AE013090 Thermoana
C 26	18	6.3	10165	1	AE006583	AE006583 Streptoco
C 27	18	6.3	13536	1	AE010066	AE010066 Streptoco
C 28	18	6.3	50463	1	AE014159	AE014159 Streptoco
29	18	6.3	348050	1	AP003581	AP003581 Nostoc sp
30	16	5.6	155932	1	CJ11168X6	AL139079 Campyloba
31	15	5.2	891	6	AX394976	AX394976 Sequence
32	15	5.2	891	6	AX398494	AX398494 Sequence
33	15	5.2	891	6	AX398804	AX398804 Sequence
34	15	5.2	4451	1	WSAJ3049	AJ003049 Wolinella
35	14	4.9	811	6	AX432600	AX432600 Sequence
36	14	4.9	894	6	AR148386	AR148386 Sequence
37	14	4.9	894	6	E05337	E05337 DNA encodin
38	14	4.9	894	6	E11866	E11866 Mutated DNA
39	14	4.9	894	6	E11867	E11867 Mutated DNA
40	14	4.9	894	6	E11869	E11869 Mutated DNA
41	14	4.9	894	6	E14762	E14762 gDNA encodi
42	14	4.9	894	6	E16095	E16095 DNA encodi
43	14	4.9	894	6	E27521	E27521 Geranyl dip
44	14	4.9	894	6	E54864	E54864 Process for
45	14	4.9	1260	1	BACFDPS	D13293 B. stearothe

ALIGNMENTS

RESULT 1

whole genome sequencing of meticillin-resistant *Staphylococcus aureus*
JOURNAL Lancet 357 (9264), 1225-1240 (2001)
MEDLINE 21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 301550)
AUTHORS Aoki,K., Oguchi,A., Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K.
and Kikuchi,H.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2001) Akio Oguchi, National Institute of
Technology and Evaluation, Biotechnology Center; 2Chome 49-10
Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:oguchienite.go.jp, URL:http://www.bio.nite.go.jp/,
Tel:81-3-3481-8423, Fax:81-3-3481-8424)
COMMENT On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES
Location/Qualifiers
source 1..301550
/organism="Staphylococcus aureus subsp. aureus N315"
/strain="N315"
/sub_species="aureus N315"
/db_xref="taxon:158879"
complement(249..947)
/gene="SA1293"
complement(249..947)
/gene="SA1293"
/note="ORFID:SA1293"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42553.1"
/db_xref="GI:13701259"
/translation="MSIFTMIIYFVILVPLWAQHKVSNYKYSQVRSSTSGKTRG
EVALTEILHANGIVDVEVGEGLTDHYDPKKVSVLSANTYDRPSVAGTAIAAHEVG
HAIQDHQGWFLFRALVPVANGSSLSYMIIMGLIILTAIGSARGSLTAGIGLM
SLAVLSIVTLPEPDPASSAMQITAINIVNEKEYKHAKVLSAAMTYVAATAVAV
AELVRILILARSDD"
complement(1001..1588)
/gene="SA1294"
complement(1001..1588)
/gene="SA1294"
/note="ORFID:SA1294"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42554.1"
/db_xref="GI:13701260"
/translation="MTINTFWOYTLQORSKMLMLICNLILMIYGIWYGEQLSHTPW
HKFIFVDPSTAILIVSILSIQKNSIIEALAFVTLFKYGIWAVIMNLILFIEQ
GDITVGLVLFMSHIMAVQAIYFPRFRSMIGISVAMIWVFLNDIYDFHLPQFPY
DFITTHWQIGVLSCLSVFGLLLYIELNKLKCK"
complement(1578..2153)
/gene="SA1295"
complement(1578..2153)
/gene="SA1295"
/note="ORFID:SA1295"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42555.1"
/db_xref="GI:13701261"
/translation="MSETLNQIKESFIEYLLFYRKPKSIAVWVLNYIKVNEAKLANI
HFVDTKINHETLAEVGSASHASAIQFTKRNKLMTNIEFDYIANHNCADFQIHFA
NYSKRQRDLQTLVAQTESPSQTYLHDLNSMAIDRKHALLIDYLLHLDLSLQMN
EKORFYOLTQILNTLKNVKNQFEDLADD"
complement(2167..3411)
/gene="SA1296"
complement(2167..3411)
/gene="SA1296"
/note="ORFID:SA1296"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB42556.1"

/db_xref="GI:13701262"
/translation="MEDIVKLIIDINLQKLENLDSRVNEAITTDNDALFILGETLYN
FGLMPOGLEVRVLYHKYPDESELLIVFIEGLMSENOTDEALEVLSVSPSPKLMLE
ADLYQQINMEVAIDKLOALELEPNPDIHFALAEMLYDGOYLRAISYEVLELG
EYQVGNLFMRMADCSLQSGNSYSDAIRLIDNEDEMTSEYDKKAISDKNDITQE
AIKIMTLTSLKDPDYIOGYLYQSLYENENKYNPDIAETGEGURLSQFYKELMYTTC
LETEHDANEVLLKQALEVDNAYQEPPLILSDLYRNEEDYEAITELLTYVDEEDLD
PTFMHLAFAYGOEERDKAQHFEAYPTMTKTNIDFMSDYFYFLIEIGQKEQAITIL
NOLLELSPSNENNHDELSQLS"
complement(3418..4716)
/gene="aroA"
complement(3418..4716)
/gene="aroA"
/note="ORFID:SA1297"
/codon_start=1
/transl_table=11
/product="3-phosphoshikimate 1-carboxyvinyltransferase"
/protein_id="BAB42557.1"
/db_xref="GI:13701263"
/translation="MVSEOIDISGLKGEIEVPGDKSMTHRAIMLASLAEGTSNIVK
PLLGDCRRTMDIFRLGVDIKEDDKLVNPSGYKAFKTPHQVLYTGNSTTRLLA
GLLSGLGIESVLGSDVSGIKRPMDRVRLPKLMDANIEGIEDNYTLIIKPSVIGKIN
YQMEASAQVKSAILFASLFSNDTIVIKELDVSRNHTETMFRHFNPIEAERUSITTT
PDAIQHKPADPHTPGDISAAPFIVAALITIPESDVTIHNVGINPTRSGIIDIVERKM
GNIQLENOTGAEPTRASIRIQYTPMLQPIITIEGELVPKAIDELPTALLCTQAVGST
IKDAEELKVKTNRNIDTADMLNLGLFELQPTNDGLIHPSEKTNATVDSLTDHRRIG
MMLAVASLSSEPVKIKQFDVNVNVPFGLPKLLENEG"
complement(4726..5790)
/gene="arob"
complement(4726..5790)
/gene="arob"
/note="ORFID:SA1298"
/codon_start=1
/transl_table=11
/product="3-dehydroquinate synthase"
/protein_id="BAB42558.1"
/db_xref="GI:13701264"
/translation="MKLOITPYSNNYPIFEHGAIDHISTYIDQFQSQFILDEHVNO
YFADKFDLILSYENVHVKLIIPAGEKTKTEQOQTELEYILSHHVTRNTALIAVGGAT
GDFAGEVAATLLRGVHFIOVPTILAHDSVGVGKVGINSKQGNLIGACYRPPAVIYD
LDFLKLPEQLSGYAEYKHALNGESTQIEQHFDRKDELQSLNGMDKRTIAGKI
ETKLDIVADEKEQGVKRNKLNHTGFEHAYEVNHNKTAHGHAVMIGIYIYFIVANILFN
SNHDIOHYINLYTKGLPLETITDIDFTIYQVYMLSDKKNDKQGVQVQVLIHKFGDILV
QHIDQTTLQACEQLKATYK"
complement(5816..6982)
/gene="aroc"
complement(5816..6982)
/gene="aroc"
/note="ORFID:SA1299"
/codon_start=1
/transl_table=11
/product="chorismate synthase"
/protein_id="BAB42559.1"
/db_xref="GI:13701265"
/translation="MRYLTSGESHGQLTIVIVEGIPANLEIKVEDINKEMFRKQGGY
RGRMQIEKDTVEIVSGVRNGYTLGSPITWVTNDQFTHWRKIMGAAPISEERENMK
RTITKPRPGHADLVGGMKYNHRDLNRVLSARSARETAARAVGALCKVLQQLDIDYI
SRVVEIGGIKDKDFDSETEKANLDNRDVIDDSTAQMRDKIDAKNKGSGISGVY
QVVENMPGVGVSVYHDKKQKIQGVVSINAFKGVSGFGEKAAERKPGSEIQDEI
LYRSDSICVAPAAISIVCEHVVAFEIAKALLEEFSQNSHIEQLQQQIADRRQLNVF"
complement(7368..7568)
/gene="SA1300"
complement(7368..7568)
/gene="SA1300"
/note="ORFID:SA1300"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB42560.1"
/db_xref="GI:13701266"
/translation="MKNIIISIIIGILMFLKMLLELYGAIFLDKPLNPITKIIFILFI
YIFYVLVKELIIFLKSYNKSA"

Db	63771	TGAACCTATTTC	CAAGTGATGATAGATAA	CTGATGAAGTAA	AAAAATAAAGTTCTTACAACG	63711
Qy	140	qLeuSerIleAla	serGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGln	ser	160	
Db	63711	GCTGTCAATAGC	CAAGTGGTCATGTTGAATG	TGCGGGTCA	AAATGTTAGATATGCAAG	63652
Qy	160	rGluGlyGlnPro	IleAspLeuGlnThrLeuGluMetIleHisLysThrLysThrGlyAl	180		
Db	63651	CGAAGGCCAACCA	ATTTGATCTTGA	AACTTTGGAAATGATACACA	AAAAACAAACAGGAGC	63592
Qy	180	aLeuLeuThrPhe	AlaValMetSerIleAlaAspIleAlaAsnValAspAspThr	ThrL	200	
Db	63591	ATTATTAACTTT	TGCGGTTTATGAGTCACGACAGATATCGCTAATGTCGATGTC	-AACTA	63533	
Qy	200	ysGluHisLeuGlu	SerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspL	220		
Db	63532	AAGACATTTAG	AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAGATGAT	63473		
Qy	220	eLeuAspCys	TyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluA	240		
Db	63472	TATTAGACTGCT	ATGCTGATGAAGCAAGTTAGTAAAAAAGTGGCAGCGATCTTGAA	63413		
Qy	240	snAsnLysSer	ThrTyrValSerLeuGlyLysAspGlyAlaGluAspLysLeuThr	260		
Db	63412	ATAATAAAGTAC	CGTACGTGAGTTTATAGGAAAGATGGCCAGCAAGATAAATGACTT	63353		
Qy	260	yrHisArgAsp	AlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysH	280		
Db	63352	ATCATGAGACG	CAGCGATGGATGAACCAATTGATGAACAATTCATACATAA	63293		
Qy	280	isLeuLeuGlu	IleValAspLeu 287			
Db	63292	ACTTATTAGAA	TCGTTGATTTA 63270			
RESULT 3						
LOCUS	AP003362/c					
DEFINITION	Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete sequence, section 5/9.					
ACCESSION	AP003362					
VERSION	BA000017					
KEYWORDS	AP003362.2 GI:14247083					
SOURCE	Staphylococcus aureus subsp. aureus Mu50 (sub_species:aureus Mu50, strain:Mu50) DNA.					
ORGANISM	Staphylococcus aureus subsp. aureus Mu50					
REFERENCE	Bacteria; Firmicutes; Bacilliales; Staphylococcus.					
AUTHORS	Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani-U, Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kunara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.					
TITLE	Whole genome sequencing of methicillin-resistant Staphylococcus aureus					
JOURNAL	Lancet 357 (9264), 1225-1240 (2001)					
MEDLINE	21311952					
REFERENCE	2 (bases 1 to 346900)					
AUTHORS	Ohta, T.					
TITLE	Direct Submission					
JOURNAL	Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College of Medical Technology and Nursing, Department of Medical Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan (E-mail:tohta@tsukuba.ac.jp. Tel:81-298-53-3454, Fax:81-298-53-3454)					
COMMENT	On May 29, 2001 this sequence version replaced gi:13875626.					
FEATURES	Location/Qualifiers					
source	1. .346900					
	/organism="Staphylococcus aureus subsp. aureus Mu50"					
	/strain="Mu50"					
	/sub_species="aureus Mu50"					
	/db_xref="taxon:158878"					


```
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="MNISGNTVTLIAVIIIGIFLQIFKPLIIVTAVLSILLGI
FVGFTVLIIVSYFNKRN"
6672..9205
/gene="nuc"
6672..9205
/gene="nuc"
/feature="SAV1324"
/codon_start=1
/transl_table=11
/product="thermonuclease"
/protein_id="BAB57486.1"
/db_xref="GI:14247095"
/translation="MKSNSLAMIIVVLIIVGVLAQFPMNHTGPKKTNHETVQDLN
```

```
Alignment Scores:
Pred. No.: 1.2e-126 Length: 346900
Score: 132.00 Matches: 285
Percent Similarity: 98.62% Conservative: 0
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 45.93% Indels: 4
DB: 1 Gaps: 0

US-09-925-637-64 (1-287) x AP003362 (1-346900)
QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
DB 251159 ATGAGCAATCTACCGATGCAATTAATATAGATGAAGTCAATTAATGATTCGGTTCG 251100
QY 21 IleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAla 40
DB 251099 ATAAATAAATCAGTAATGATGATCTCAGCTAGAGAAAGCATGCTGTTATTCATTAATGCT 251040
QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 251039 GGAGTAAACGCATCCGACCATCTCTGTTATCTACTCTTATGATTCATTAATACCGAG 250980
QY 61 TyrGluLeuGlyMet-LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeu1 80
DB 250979 TATGAGTTAGGTGT-GAAGAGCGCAATTCGACTAGAAATGATTCATATTCATTCAT 250921
QY 80 eHisAspAspLeuProAlaMetAspAspTyrArgArgGlyLysLeuThrAsnH1 100
DB 250920 TCATGATGACCTACCGAGTGGATAATGATGATGATGATGATGATGATGATGATGATGAT 250861
QY 100 sLysValTyrGlyGluTyrThrAlaIleLeuAlaGlyAspAlaLeuLeuLeuLeuLeu 120
DB 250860 TAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250801
QY 120 eGluLeuLeuSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnAr 140
DB 250800 TGAACCTATTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250741
QY 140 gLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnse 160
DB 250740 GCTGTCAATAGCAAGTGTGATGTTGGAATGCTGCGCGTCAATGTTAGATATGCAAG 250681
QY 160 cLysGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAl 180
DB 250680 CGAAGGCCAACCAATGATGTTGAACTTGGAAATGATACACAAACAAACAGGAGC 250621
QY 180 aLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-ThrL 200
DB 250620 ATTATTAATTTTTCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250562
QY 200 ysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspL 220
DB 250561 AAGAACAATTTAGAAAGTTATAGTTATCATTTAGGTATGATGATGATGATGATGATGATGAT 250502
QY 220 euLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluA 240
```

```
DB 250501 TATTAGACTGCTATGATGAAGCAAGTTAGGTAAAGAGTGGCAGCATCTGAAA 250442
QY 240 snAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrT 260
DB 250441 ATAATAAAAGTACGTACGTGAGTTTATAGGAAAGATGGCCAGAGATAAATTCAC 250382
QY 260 yHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysH 280
DB 250381 ATCATAGACGACGACGAGTGGATGAATCAACCAATTCATGAACAATTCATAACAA 250322
QY 280 sLeuLeuGluLeuValAspLeu 287
DB 250321 ACTTATTAGAAATCGTTGATTTA 250299
RESULT 4
LOCUS AR106466 882 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6107058.
ACCESSION AR106466
VERSION AR106466.1 GI:12820996
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 882)
AUTHORS Gwynn.M. and Wilding, E.Imogen.
TITLE isPA from Staphylococcus aureus
JOURNAL Patent: US 6107058-A 1 22-AUG-2000;
FEATURES
source
BASE COUNT 328 a 118 c 186 g 250 t
ORIGIN
Alignment Scores:
Pred. No.: 1.91e-97 Length: 882
Score: 102.00 Matches: 284
Percent Similarity: 97.93% Conservative: 0
Best Local Similarity: 97.93% Mismatches: 3
Query Match: 35.54% Indels: 6
DB: 6 Gaps: 0

US-09-925-637-64 (1-287) x AR106466 (1-882)
QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
DB 1 ATGAGCAATCTACCGATGCAATTAATATAGATGAAGTCAATTAATGATTCGGTTCG 60
QY 21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
DB 61 ATAAATAAATCAGTAATGATGATCTCAGCTAGAGAAAGCATGCTGTTATTCATTAATGCT 120
QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 121 GGAGTAAACGCATCCGACCATCTCTGTTATCTACTCTTATGATTCATTAATACCGAG 180
QY 61 TyrGluLeuGlyMet-LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeu1 80
DB 181 TATGAGTTAGGTGT-GAAGAGCGCAATTCGACTAGAAATGATTCATATTCATTCAT 239
QY 80 eHisAspAspLeuProAlaMetAspAspTyrArgArgGly-LysLeuThrAsnH 100
DB 240 TCATGATGACCTACCGAGTGGATAATGATGATGATGATGATGATGATGATGATGATGAT 298
QY 100 sLysValTyrGlyGluTyrThrAlaIleLeuAlaGlyAspAlaLeuLeuLeuLeuLysAlaP 120
DB 299 ATAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 358
QY 120 heGluLeuLeuSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnA 140
DB 359 TTGAACCTATTTCAAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418
```

QY 140 rgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlns 160
Db 419 GCGTGTCAATAGCAAGTGGTCATGTTGGAATGTCGCGGTCAATGTTAGATATGCAAA 478
QY 160 erGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyA 180
Db 479 GCGAAGGCCCAACCAATTCATGCTTGAACCTTGGAAATGATACAAACAAACAGGAG 538
QY 180 laLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspThr-Thr 199
Db 539 CATATTAACTTTTCGCTTATGATGTCAGCAGATATCGCTAATGTCGATGTC-AACT 597
QY 200 LysGluHisLeuGluSerTyrrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
Db 598 AAGAACAATTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAGATGAT 657
QY 220 LeuLeuAspCysTyrrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGlu 239
Db 658 TTATTAGACGCTATGCTGATGATGAAGCAAAAGTTAGTTAAAAAGTGGCCAGCATCTTGAA 717
QY 240 AsnAsnLysSerThrTyrrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
Db 718 AATAATAAAGTAGTACGTAGTGTATATTAGGGAAGATGCGCAGAGATAAATGACT 777
QY 260 TyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLys 279
Db 778 TATCATAGACGACGACGAGTGGATGAACCAATGATGAACAATTCATACATAA 837
QY 280 HisLeuLeuGluIleValAspLeu 287
Db 838 CACTTATTAGAAATCGTTGATTTA 861

RESULT 5
AX141501
LOCUS AX141501 783 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 223 from Patent WO0134809.
ACCESSION AX141501
VERSION AX141501.1 GI:14281622
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 783)
AUTHORS Kimmerly, W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 223 17-MAY-2001;
GLAXO GROUP LIMITED (GB)

FEATURES
source
1..783
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
BASE COUNT 297 a 93 c 147 g 246 t
ORIGIN
Alignment Scores:
Pred. No.: 3.62e-29 Length: 783
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.89% Indels: 0
DB: 6 Gaps: 0
US-09-925-637-64 (1-287) x AX141501 (1-783)

QY 70 AlaLeuGluMetIleHisThrTyrrSerLeuIleHisAspLeuProAlaMetAspAsn 89
Db 109 GCATTGGAATGATTCATCTATTCTTTAATTCATGATGATTACCAGCAATGGATAAT 168
QY 90 AspAspTyrrArgGlyLysLeuThrAsnHisLysValTyrrGlyGluTrp 106
Db 169 GACGATTACCGTAGAGGAAAAATTACAAATCATATAAGTTATGTTGTAATGG 219

RESULT 6
AF270104
LOCUS AF270104 3444 bp DNA linear BCT 01-AUG-2000
DEFINITION Staphylococcus epidermidis strain SRI clone step.1049b08 genomic
sequence.
ACCESSION AF270104
VERSION AF270104.1 GI:9624010
KEYWORDS
SOURCE Staphylococcus epidermidis.
ORGANISM Staphylococcus epidermidis.
REFERENCE 1 (bases 1 to 3444)
AUTHORS Kimmerly, W.J., Taylor, J. David., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
Listenebee, S., Ashanti, C., Altschuller, G., Mamo, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3444)
AUTHORS Taylor, J. David., Kimmerly, W.J., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
Listenebee, S., Ashanti, C., Altschuller, G., Mamo, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA

FEATURES
source
1..3444
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1049b08"
BASE COUNT 1365 a 451 c 571 g 1057 t
ORIGIN
Alignment Scores:
Pred. No.: 1.37e-28 Length: 3444
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.89% Indels: 0
DB: 1 Gaps: 0
US-09-925-637-64 (1-287) x AF270104 (1-3444)

QY 70 AlaLeuGluMetIleHisThrTyrrSerLeuIleHisAspLeuProAlaMetAspAsn 89
Db 973 GCATTGGAATGATTCATCTATTCTTTAATTCATGATGATTACCAGCAATGGATAAT 1032
QY 90 AspAspTyrrArgGlyLysLeuThrAsnHisLysValTyrrGlyGluTrp 106
Db 1033 GACGATTACCGTAGAGGAAAAATTACAAATCATATAAGTTATGTTGTAATGG 1083
RESULT 7
AX145422
LOCUS AX145422 3444 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 4144 from Patent WO0134809.
ACCESSION AX145422
VERSION AX145422.1 GI:14283987
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 3444)
AUTHORS Kimmerly, W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 4144 17-MAY-2001;
GLAXO GROUP LIMITED (GB)

FEATURES
Location/Qualifiers

QY	70	AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProIaMetAspAsn	89
Db	150	GCATTGGAATGATTCATCTATTCTTTAAATTCATGATTTACCAGCAATGGATAAT	91
QY	90	AspAspTyrArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp	106
Db	90	GACGATTACCGTAGAGGAAATTTAACAAATCATATAAGTTTATGGTGAATGG	40
RESULT 9			
LOCUS	AX145276/c		
DEFINITION	Sequence 3998 from Patent WO0134809.		
ACCESSION	AX145276		
VERSION	AX145276.1	GI:14283841	
KEYWORDS	synthetic construct.		
SOURCE	artificial construct.		
ORGANISM	artificial sequences.		
REFERENCE	1 (bases 1 to 3454)		
AUTHORS	Kimmerly, W.J.		
TITLE	Staphylococcus epidermidis nucleic acids and proteins		
JOURNAL	Patent: WO 0134809-A 3998 17-MAY-2001;		
	GLAXO GROUP LIMITED (GB)		
FEATURES	Location/Qualifiers		
source	1..3454		
BASE COUNT	1005 a 574 c 493 g 1382 t		
ORIGIN	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="synthetic nucleic acid sequence"		
Alignment Scores:			
Pred. No.:	1.38e-28	Length:	3454
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.89%	Indels:	0
DB:	6	Gaps:	0
US-09-925-637-64 (1-287) x AX145276 (1-3454)			
QY	70	AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProIaMetAspAsn	89
Db	150	GCATTGGAATGATTCATCTATTCTTTAAATTCATGATTTACCAGCAATGGATAAT	91
QY	90	AspAspTyrArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp	106
Db	90	GACGATTACCGTAGAGGAAATTTAACAAATCATATAAGTTTATGGTGAATGG	40
RESULT 10			
LOCUS	AF269889/c		
DEFINITION	Staphylococcus epidermidis strain SRI clone step.1028f08 genomic sequence.		
ACCESSION	AF269889		
VERSION	AF269889.1	GI:9623789	
KEYWORDS	Staphylococcus epidermidis.		
SOURCE	Staphylococcus epidermidis		
ORGANISM	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
REFERENCE	1 (bases 1 to 4045)		
AUTHORS	Kimmerly, W.J., Taylor, J. David., Nelsen, A.J., Godlevski, M.M., Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I., Listenbee, S., Ashanti, C., Altshuler, G., Mamo, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Furdon, P.J.		
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis genome		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 4045)		
AUTHORS	Taylor, J. David., Kimmerly, W.J., Nelsen, A.J., Godlevski, M.M., Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,		

Listenbee, S., Ashanti, C., Altshuller, G., Mamo, L., Shepherd, N. S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D. H., Miller, G. S. and Furdon, P. J.

Direct Submission
Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA

```

FEATURES
  source
    location/Qualifiers
      1..4045
        /organism="Staphylococcus epidermidis"
        /strain="SRI"
        /db_xref="taxon:1282"
        /clone="step.1028f08"
1249 a 691 c 536 g 1569 t
BASE COUNT

```

Alignment Scores:	1.59e-28	Length:	4045
Pred. No.:	37.00	Matches:	37
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
1st Local Similarity:	100.00%	Indels:	0
Query Match:	12.89%	Gaps:	0
DB:	1		

US-09-925-637-64 (1-287) x AF269889 (1-4045)

Qy	70	AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn	89
Db	2106	GCATTGGAAATGATTTCATCTATTCTTTAATTCATGATCATTTTACCAGCAATGGATAAT	2047

Alignment Scores:		
Pred. No.:	1.59e-28	4045
Score:	37.00	37
Percent Similarity:	100.00%	
Best Local Similarity:	100.00%	
Query Match:	12.89%	
DB:	6	
Length:		
Matches:		
Conservative:		
Mismatch:		
Indels:		
Gaps:		

US-09-925-637-64 (1-287) x AX145207 (1-4045)

70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
 2106 GCATTGGAAATGATTTCATCTTATCTTTAAATTCATGATTTCACAGCAATGGATAAT 2047

Db 2046 GACGATTACCGTAGAGGAAAAATTAAACAAATCATAAAGTTTATGGTGAATGG 1996

Alignment Scores:			
Pred. No.:	3,786-25	Length:	486
Score:	33.00	Matches:	33
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.50%	Indels:	0
DB:	6	Gaps:	0

US-09-925-637-64 (1-287) x AR097529 (1-486)

Qy 69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88

Alignment Scores:	3.78e-25	486
Pred. No.:	33.00	33
Score:	100.00%	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	11.50%	Indels:
DB:	6	Gaps:
		0

US-09-925-637-64 (1-287) x AR142282 (1-486)

69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAsp 88

QY 89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
|||||
DB 145 AATGATGACTATCGTCGGGTAACACTCAGATCATAAG 183
|||||

RESULT 14
E15319
LOCUS E15319 486 bp DNA linear PAT 28-JUL-1999
DEFINITION Micrococcus luteus FPS gene fragment.
ACCESSION E15319
VERSION E15319.1 GI:5710002
KEYWORDS JP 1998057079-A/10.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 486)
AUTHORS Muramatsu,M., Koike,A., Ogura,K., Furuyama,T., Shimizu,N. and Chiyou,G.
TITLE GENE FOR PRENYL DIPHOSPHORIC ACID-SYNTHESIZING ENZYME
JOURNAL Patent: JP 1998057079-A 10 03-MAR-1998;
COMMENT TOYOTA MOTOR CORP
OS Micrococcus luteus
PN JP 1998057079-A/10
PD 03-MAR-1998
PE 13-JUN-1997 JP 1997157228
PR 14-JUN-1996 JP 96P 154441
PI MURAMATSU MASAYOSHI, KOIKE AYUMI, OGURA KYOZO, PI FURUYAMA TANETOSHI,
SHIMIZU NAOTO, CHIYOU GENI
PC C12N15/09,C07H21/04,C12N1/21,C12N9/00,C12R1/125),
PC C12N15/09,
PC C12R1/07),(C12N1/21,C12R1/19),(C12N9/00,C12R1/19); CC
strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1. 486 /organism="Micrococcus luteus".
FT location/Qualifiers
1. 486
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 148 a 85 c 98 g 149 t 6 others
ORIGIN

Alignment Scores:
Pred. No.: 3.78e-25 Length: 486
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.50% Indels: 0
DB: 6 Gaps: 0

US-09-925-637-64 (1-287) x E15319 (1-486)

QY 69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
|||||
DB 85 ATTCGGCTTGAATGATTTCATACGATTCCTTTAAATTCACGATGACTTCGCGCAATGAT 144
|||||

QY 89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
|||||
DB 145 AATGATGACTATCGTCGGGTAACACTCAGATCATAAG 183
|||||

RESULT 15
AB003187
LOCUS AB003187 1080 bp DNA linear BCT 25-MAR-1998
DEFINITION Micrococcus luteus DNA for farnesyl diphosphate synthase, complete cds.
ACCESSION AB003187
VERSION AB003187.1 GI:2982676
KEYWORDS farnesyl diphosphate synthase.
SOURCE Micrococcus luteus (strain:B-P 26) DNA.
ORGANISM Micrococcus luteus

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Micrococccineae; Micrococccaceae; Micrococcus.
1 (sites)
Shimizu,N., Koyama,T. and Ogura,K.
Molecular cloning, expression, and characterization of the genes
encoding the two essential protein components of Micrococcus luteus
B-P 26 hexaprenyl diphosphate synthase
J. Bacteriol. 180 (6), 1578-1581 (1998)
98175686
2 (bases 1 to 1080)
koyama,T.
Direct Submission
Submitted (18-APR-1997) Tanetoshi koyama, Tohoku University,
Institute for Chemical Reaction Science; Katahira 2-1-1, Aoba-ku,
Sendai, Miyagi 980-77, Japan (E-mail:koyama@icrs.tohoku.ac.jp,
Tel:+81-22-217-5621, Fax:+81-22-217-5620)
Location/Qualifiers
1. 1080
/organism="Micrococcus luteus"
/strain="B-P 26"
/db_xref="taxon:1270"
105..980
/gene="fps"
105..980
/gene="fps"
/codon_start=1
/transl_table=11
/product="farnesyl diphosphate synthase"
/protein_id="BAA25265.1"
/db_xref="GI:2982677"
/translation="MQEKLWNRDFLNLINESLLKHYHPAQSRLEHAINYSLSAGGK
RIRPLVITLDSLGGNAHDLPGFIALEMHTYSLIHDDLPAMDNDYRRKLTNHH
RFEATATLADALITDAFCILATQNAEIKLSLNLSTASGSGMVGOMLDMQG
EKKTLITNELERIHIGKELIRAAIVSAGIIMNFDAIQEQLNTIGKNVGLMFOIKD
DILDVEGSEFENIGTVGSDLNDRKSTVYVLLGLEASKOLLNDKLTETDYDAUKTLQPIN
DNLKLITYIVERNK"
BASE COUNT 377 a 161 c 198 g 344 t
ORIGIN

Alignment Scores:
Pred. No.: 7.75e-25 Length: 1080
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.50% Indels: 0
DB: 1 Gaps: 0

US-09-925-637-64 (1-287) x AB003187 (1-1080)

QY 69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
|||||
DB 312 ATTCGGCTTGAATGATTTCATACGATTCCTTTAAATTCACGATGACTTCGCGCAATGAT 371
|||||

QY 89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
|||||
DB 372 AATGATGACTATCGTCGGGTAACACTCAGATCATAAG 410
|||||

Search completed: May 30, 2003, 23:41:52
Job time : 2998 secs

30-JUL-1997.

07-JAN-1997; 97EP-0100117.

05-JAN-1996; 96US-0009861.

(HUMA-) HUMAN GENOME SCI INC.

Barash SC, Choi GH, Dillion PJ, Fannon MR, Kunsch CA;

Rosen CA;

WPI; 1997-374922/35.

Poly-nucleotide(s) and proteins derived from *Staphylococcus aureus* stored on computer readable medium and used in the production of anti-S. aureus vaccines

Claim 1; Page 784-785; 3271pp; English.

This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S. aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S. aureus infection. The polypeptides can also be used in a kit for the immunodetection of S. aureus in a sample. S. aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the S. aureus DNA sequences contained on the computer readable medium.

Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

Alignment Scores:

Pred. No.: 9,02e-262 Length: 1893
Score: 269.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.73% Indels: 0
DB: 18 Gaps: 0

us-09-925-637-64 (1-287) x AAV74466 (1-1893)

19 ValAlaIleAsnLysSerValMetAspThrGluGluSerMetLeuTyrSerLeu 38
1260 GTTGGGATAAATAAATCAGTAATGGATACCTCAGCTAGAGAAAGTATGTTGATTCATTA 1201
39 AsnAlaGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 58
1200 AATGCTGGAGTAAACCATCCGACCATCTGTTGTTATCTACTCTTTAGATTCCTAATAAT 1141
59 ThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSer 78
1140 ACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
79 LeuIleHisAspLeuProAlaMetAspAspTyrArgGlyLysLeuThr 98
1080 CTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
99 AsnHisLysValTyrGlyGluThrThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLys 118
1020 AATCATAAGTATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961
119 AlaPheGluLeuIleSerSerAspArgLeuThrAspGluValLysIleLysValLeu 138
960 GCATTTGAACCTTATTTCAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901

QY 139 GlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMet 158
DB 900 CAACGGCTGTCAATAGCAAGTGGTCATCTTGAATGGTGGCGGTCAAAATGTTAGATATG 841
QY 159 GlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThr 178
DB 840 CAAAGCGAAGGCCAACCAATTTGATCTTGAATTTGGAAATGATACACAAACAAAACA 781
QY 179 GlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr 198
DB 780 GGAGCATTTATTAACCTTTGGGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 721
QY 199 ThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAsp 218
DB 720 ACTAAAGAACAATTTAGAAAGTTATATGATCTTATGATGATGATGATGATGATGATGATGATGAT 661
QY 219 AspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeu 238
DB 660 GATTATTAGATGCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
QY 239 GluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeu 258
DB 600 GAAATATAAAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 541
QY 259 ThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThr 278
DB 540 ACTTATCATAGAGACGACGACGATGGATGAACCAACTAAGCAAAATTTGATGAACATTCATACA 481
QY 279 LysHisLeuLeuGluIleValAspLeu 287
DB 480 AACACTTATTAGAAATCGTTGATTTA 454
RESULT 3
AAS54848
ID AAS54848 standard; DNA; 882 BP.
XX AAS54848;
AC AAS54848;
XX 13-FEB-2002 (first entry)
XX Staphylococcus aureus DNA for cellular proliferation protein #1160.
DE Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX Staphylococcus aureus.
OS WO200170955-A2.
PN 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US09180.
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207272P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI WPI; 2001-611495/70.
DR P-PSDB; AAU36989.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX

181 AAGAGCCGATTCGACCTAGAAATGATTCATACATATTCATCTATTCATGATGACCTACCA 240
256 GCGATGATTAATGATGATTCGACGAGGAGAAATTAACAAATCATTAAGTATATGTTGAG 315
241 GCGATGATTAATGATGATTCGACGAGGAGAAATTAACAAATCATTAAGTATATGTTGAG 300
316 TGGACTGCGATATAGCAGGTGATGCTTTATTAACATAACATTTGAACATTTTCAAGT 375
301 TGGACTGCGATATAGCAGGTGATGCTTTATTAACATAACATTTGAACATTTTCAAGT 360
376 GATGATAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
361 GATGATAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
436 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
421 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
496 GATCTTGAACCTTTGGAAATGATACACAAACAAACAAACAAACAAACAAACAAACAAACAA 555
481 GATCTTGAACCTTTGGAAATGATACACAAACAAACAAACAAACAAACAAACAAACAAACAA 540
556 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 615
541 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
616 TATAGTTATCATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
601 TATAGTTATCATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
676 GAT 735
661 GAT 720
736 GTGAT 795
721 GTGAT 780
796 GTGAT 855
781 GTGAT 840
856 GATTTA 861
841 GATTTA 846

RESULT 5
V74466/c
AAV74466 standard; DNA; 1893 BP.
AAV74466;
16-MAR-1999 (first entry)
Staphylococcus aureus contig SEQ ID #155.
Computer readable medium; vaccine; S.aureus infection; immunodetection;
cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
skin infection; surgical wound infection; scalded skin syndrome;
toxic shock syndrome; ds.
Staphylococcus aureus.
Key Location/Qualifiers
misc_feature 1261..1320
/tag-a
/note- these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence.

PN EP786519-A2.
XX 30-JUL-1997.
XX 07-JAN-1997; 97EP-0100117.
XX 05-JAN-1996; 96US-0009861.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
XX Rosen CA;
XX WPI; 1997-374922/35.
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus
XX stored on computer readable medium and used in the production of
XX anti-S.aureus vaccines
XX Claim 1; Page 784-785; 3271pp; English.
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
XX medium, preferably selected from a floppy or hard disk, random access
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX the S.aureus DNA sequences allows putative functions to be assigned so
XX that protein-encoding or regulatory regions of commercial, therapeutic or
XX industrial importance can be obtained. Specifically, sequences which are
XX likely to encode antigens have been identified and these polypeptides can
XX be used in a vaccine composition against S.aureus infection. The
XX polypeptides can also be used in a kit for the immunodetection of
XX S.aureus in a sample. S.aureus is implicated in numerous human diseases,
XX including cellulitis, eyelid infections, food poisoning, osteomyelitis,
XX skin and surgical wound infections, scalded skin syndrome, toxic shock
XX syndrome, etc. Organisms transformed with the DNA sequences can be used
XX for recombinant production of the polypeptides. The new DNA sequences
XX (and their fragments) are useful as primers or probes for isolating
XX homologues of any of the S.aureus DNA sequences contained on the
XX computer readable medium.
XX Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;
Query Match 93.7%; Score 807; DB 18; Length 1893;
Best Local Similarity 100.0%; Pred. No. 2.5e-181;
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 GTTGGATTAATAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 114
DB 1260 GTTGGATTAATAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12
QY 115 AATGCTGGAGGTAAGCGATCCGACGAGTCTCTGTTATTAATCACTTTAGATTCTACTAAAT 174
DB 1200 AATGCTGGAGGTAAGCGATCCGACGAGTCTCTGTTATTAATCACTTTAGATTCTACTAAAT 1141
QY 175 ACCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 234
DB 1140 ACCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
QY 235 CTTATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 294
DB 1080 CTTATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
QY 295 AATCAATAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
DB 1020 AATCAATAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961
QY 355 GCATTTGAACCTTTATTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 414
DB 960 GCATTTGAACCTTTATTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
QY 415 CAACGCTGTCATAGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 474
DB 900 CAACGCTGTCATAGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 841

QY	475	CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAACAAAAACA	534
Db	840	CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAACAAAAACA	781
QY	535	GGAGCATTAATTAACCTTTTCGGTTATGAGTCCACGAGATATCGGTATGTCGATGATACA	594
Db	780	GGAGCATTAATTAACCTTTTCGGTTATGAGTCCACGAGATATCGGTATGTCGATGATACA	721
QY	595	ACTAAAGAACAATTTAGAAGTTATAGTTATCATTTAGTGTATGATGTTCCAGATTTAAAGAT	654
Db	720	ACTAAAGAACAATTTAGAAGTTATAGTTATCATTTAGTGTATGATGTTCCAGATTTAAAGAT	661
QY	655	GATTTATTAGACTGCTATGGTGATGAAGCAAGTTAGGTAAAAAGTTGGCAGCGATCTT	714
Db	660	GATTTATTAGACTGCTATGGTGATGAAGCAAGTTAGGTAAAAAGTTGGCAGCGATCTT	601
QY	715	GAAATAATAAAGTACGTACGTGAGTTTATTACGGAAAGATGGCGCAGAAGATAAATG	774
QY	600	GAAATAATAAAGTACGTACGTGAGTTTATTACGGAAAGATGGCGCAGAAGATAAATG	541
QY	775	ACTTATCATAGACGCGACGAGTGGATGAACCTAACGCCAAATTTGATGAACAATTTCAATACA	834
Db	540	ACTTATCATAGACGCGACGAGTGGATGAACCTAACGCCAAATTTGATGAACAATTTCAATACA	481
QY	835	AAACACTTATTAAGAATCGTTGATTTA	861
Db	480	AAACACTTATTAAGAATCGTTGATTTA	454

RESULT 6

ABN90871

ID ABN90871 standard; DNA: 909 BP.

XX

AC ABN90871;

XX

DT 24-JUL-2010

XX

DE Staphylococcus aureus

XX

KW **Staphylococcus**

KW antibacte.

XX
C+amblyo

US Staphylococcus

XX
DN
PC6390370

PN . US6380370
YY

30-APR-2011

30-APR-20

13-AUG-1964

XX
XX
CT NOV 27

PR 14-AUG-19

PR 08-NOV-19

XX XX

PA (GENO-) GEN

XX

PI Doucette-

DR WPI; 2002

CC cycle or inhibit *S. epidermidis* infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.

RESULT 7

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 21:50:41 ; Search time 276 Seconds
(without alignments)
2341.752 Million cell updates/sec

Title: US-09-925-637-64

Perfect score: 287

Sequence: 1 MTNLPNNKLIDEVNNELSA.....ELTQIDRQFNFKHLLLEIVDL 287

Scoring table:

ORIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Arched: 2185239 seqs, 1125999159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4368727

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO_spool/US09925637/runat_23052003.174850_24102/app_query.fasta.1.455
-DB=N_Geneseq_101002 -OFT=Fastap -SUFFIX=oli.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=colligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637 -CGN_1_1_263 -runat_23052003.174850_24102 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	287	100.0	861	22	AAS00821	S. aureus HGS072 e
2	269	93.7	1893	18	AAV74466	Staphylococcus aur
3	197	68.6	882	23	AAS54848	Staphylococcus aur
4	192	66.9	864	23	AAS51602	Staphylococcus aur
5	137	47.7	413	23	AAS50245	Staphylococcus aur
6	112	39.0	337	23	AAS50698	Staphylococcus aur
7	102	35.5	882	21	AAA92031	Staphylococcus aur
8	37	12.9	783	22	AAH52415	S. epidermidis ope
9	37	12.9	909	24	ABN90871	Staphylococcus epi
10	37	12.9	3444	22	AAH54780	S. epidermidis gen
11	37	12.9	3454	22	AAH54634	S. epidermidis gen
12	37	12.9	4045	22	AAH54565	S. epidermidis gen
13	33	11.5	486	19	AAV15824	Probe.B500 for Mic
14	27	9.4	5635	24	ABQ71010	Listeria monocytog
15	27	9.4	2944528	24	ABA03041	Listeria innocua c
16	19	6.6	495269	24	ABQ67195	Listeria innocua c
17	19	6.6	3011208	24	ABQ69245	Listeria innocua D
18	18	6.3	870	24	ABN66185	Streptococcus poly
19	18	6.3	882	23	AAS52896	Enterococcus faeca
20	18	6.3	7528	20	AAI12992	Enterococcus faeca
21	18	6.3	2365589	24	ABA90521	Genomic sequence o
22	15	5.2	891	24	ABK83265	High growth methan
23	15	5.2	891	24	AAD35506	Methylomonas 16a s
24	15	5.2	891	24	ABK50087	Methylomonas 16a O
25	14	4.9	811	24	ABK73724	Bacillus lichenifo
26	14	4.9	873	24	ABN66184	Streptococcus poly
27	14	4.9	894	14	AAQ39243	FPS DNA. Bacillus
28	14	4.9	894	17	AAQ40226	Mutant farnesylidip
29	14	4.9	894	17	AAQ40227	Mutant farnesylidip
30	14	4.9	894	17	AAQ40229	Native farnesylidip
31	14	4.9	894	19	AAV38455	DNA encoding farne
32	14	4.9	894	19	AAV18516	Bacillus stearothe
33	14	4.9	894	20	AAH86781	Farnesyl diphospha
34	14	4.9	894	21	AAC63809	Bacillus stearothe
35	14	4.9	894	24	ABK96797	B. stearothermophi
36	14	4.9	2155561	24	ABN71527	Streptococcus poly
37	12	4.2	888	23	AAS53505	Haemophilus influe
38	12	4.2	897	20	AAZ33160	Escherichia coli n
39	12	4.2	900	23	AAS52335	E. coli DNA for ce
40	12	4.2	900	23	AAS56349	Salmonella typhi D
41	12	4.2	900	24	AAI40785	DNA relating to th
42	12	4.2	900	24	AAI40785	E.coli prenyl dip
43	12	4.2	903	24	ABQ90268	M. capsulatus gene
44	12	4.2	4390	20	AAZ33182	Escherichia coli D
45	12	4.2	11692	22	AA546249	DNA encoding novel

ALIGNMENTS

RESULT 1

AAS00821

ID AAS00821 standard; DNA; 861 BP.

XX AC AAS00821;

XX DF 04-JUL-2001 (first entry)

XX DE AAS00821 encoding Farnesyl diphosphatesynthase, Ispa.

XX KW Farnesyl diphosphatesynthase; Ispa; immunogen; vaccine; antibody;

XX KW wound infection; cellulitis; burn infection; eyelid infection;

XX KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;

XX KW skin infection; scalded skin syndrome; toxic epidermal necrosis;

XX KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;

XX KW HGS072; ds.

XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers
 FT CDS 1..861
 FT /*tag= a
 FT /*product= "IspA"
 FT /*partial
 FT /*note= "No stop codon"
 XX W0200116292-A2.
 XX 08-MAR-2001.
 XX 31-AUG-2000; 2000WO-US23773.
 XX 01-SEP-1999; 99US-0151933.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX ChOI GH;
 XX WPI; 2001-183259/18.
 XX P-PSDB; AAU00859.
 XX New isolated nucleic acid for use in diagnosing Staphylococcus
 PT infections and in vaccines for eliciting immune responses to the
 PT infections
 XX
 XX Claim 1; Page 23; 225pp; English.
 XX The sequence encodes S. aureus IspA (Farnesyl diphosphatesynthase).
 CC The polynucleotides of the invention are used to detect Staphylococcus
 CC nucleic acids in a biological sample from an animal for diagnosing
 CC Staphylococcus infections. The polypeptides of the invention are used to
 CC detect anti-Staphylococcus antibodies in a biological sample from an
 CC animal to diagnose Staphylococcus infections. The polypeptides are also
 CC used in vaccines to elicit protective antibodies in an animal to a member
 CC of the Staphylococcus genus and for preventing or attenuating an
 CC infection caused by a member of the Staphylococcus genus e.g wound
 CC infection, cellulitis, burn infection, eyelid infection, food poisoning,
 CC joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
 CC scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
 CC disease and Lyell's disease), toxic shock syndrome and endocarditis. The
 CC polynucleotides may also be used in vaccines and for preventing or
 CC attenuating a Staphylococcus infection. Antibodies to the polypeptides
 CC may be used to purify, detect and target the polypeptides in vitro and
 CC in vivo diagnostic and therapeutic methods.
 XX
 XX Sequence 861 BP; 373 A; 114 C; 180 G; 244 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,79e-280 Length: 861
 Score: 287.00 Matches: 287
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps:
 US-09-925-637-64 (1-287) x AAS00821 (1-861)
 QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
 DB 1 ATGACGAATCTACCGATGAATAAATAATAGTAGAAGTCAATAATGATTCGGTGGC 60
 QY 21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
 DB 61 ATAATAATCATCATGATGATCTACAGTACAGTAGAAGAATGTTGTTATCATTAATGCT 120
 QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 DB 121 GCAGGTAACCGCATCCGACGAGTCTGTTATTACTCATTAGATTCACTAAATACCGAG 180
 QY 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80

DB 181 TATGAGTTAGGTATGAAGAGCGCAATGTCACCTAGAAATGATTATCATATATTCACCTATT 240
 QY 81 HisAspAspLeuProAlaMetAspAsnAspTyrArgArgGlyLysLeuThrAsnHis 100
 DB 241 CATGATGACCTACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
 DB 301 AAAGTATATGCTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
 QY 121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
 DB 361 GAACCTTATTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyClnMetLeuAspMetGlnSer 160
 DB 421 CTGTCAATAGCAAGTGTCTGTTGGAAATGCTCGCGGTCAAAATGTTAGATATGCAAGC 480
 QY 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
 DB 481 GAAGGCCAACCAATTGATCTTGAAATCTTGGAATGATACACAAACAAACACAGAGCA 540
 QY 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrLys 200
 DB 541 TTATTAACTTTTTCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
 DB 601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAGATGATTTA 660
 QY 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsn 240
 DB 661 TTAGACTGCTATGCTGATGAAGCAAAAGTTAGGTAAAGTGGCAGCATCTTGAAT 720
 QY 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaLysLeuThrTyr 260
 DB 721 AATAAAGTACGTACGTACGTAGTTATTAGGAAAGATGGCGCAGAGATTAATTCATTTAT 780
 QY 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
 DB 781 CATAGAGCGCAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 QY 281 LeuLeuGluIleValAspLeu 287
 DB 841 TTATTAGAAATCGTTGATTTA 861
 RESULT 2
 ID AAV74466 standard; DNA; 1893 BP.
 XX AC AAV74466;
 XX AC
 DT 16-MAR-1999 (first entry)
 XX
 DE Staphylococcus aureus contig SEQ ID #155.
 XX
 KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 1261..1320
 FT /*tag= a
 FT /*note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 XX
 PN EP786519-A2.
 XX

PD 30-JUL-1997.
XX
PF 07-JAN-1997; 97EP-0100117.
XX
PR 05-JAN-1996; 96US-0009861. ✓
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX
DR WPI; 1997-374922/35.
XX
XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus*
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
PS Claim 1; Page 784-785; 3271pp; English.
XX
XX This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
XX
SQ Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

Alignment Scores:
Pred. No.: 9.02e-262 Length: 1893
Score: 269.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.73% Indels: 0
DB: 18 Gaps: 0

us-09-925-637-64 (1-287) x AAV74466 (1-1893)

19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeu 38
1260 GTTCGGATAATAATCAGTAATGGATCTCAGCTAGAGAAAGTATGTTGATTCATTA 1201

39 AsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsn 58
1200 AATGCTGGAGTAACCAACCACTCCGACCACTTCGTATTATCTACCTTTAGATTCATAAT 1141

59 ThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSer 78
1140 ACCGAGTATGATGTAGTATGATGAGAGCGCAATTCCTAGAAATGATTCATACATATCA 1081

79 LeuIleHisAspLeuProAlaMetAspAsnAspTyrArgArgGlyLysLeuThr 98
1080 CTTATTCATGATGACCTACCGACCATGATGATGATGATGATGATGATGATGATGATGAT 1021

99 AsnHisLysValTyrGlyGluThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLys 118
1020 AATCATAAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961

119 AlaPheGluLeuIleSerSerAspArgLeuThrAspGluValLysIleLysValLeu 138
960 GCATTTGAACTATTTCAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901

Oy 139 GlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMet 158
Db 900 CAACGGCTGTCAATAGCAAGTGGTCATGTTGAAATGGTGGCGGTCAATGTTAGATATG 841

Oy 159 GlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThr 178
Db 840 CAAAGCGAAGCCCAACCAATTTGATCTTTGAAACTTTTGGAAATGATACACAAACAAACA 781

Oy 179 GlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr 198
Db 780 GGACATTTATTAACCTTTTCGGTTATGATGATGATGATGATGATGATGATGATGATGAT 721

Oy 199 ThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAsp 218
Db 720 ACTAAGAACAATTTAGAAAGTTATAGTTATCATTTAGTATGATGATGATGATGATGATGAT 661

Oy 219 AspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlyLysValGlySerAspLeu 238
Db 660 GATTTATTTAGACTGCTATGTTGATGAAGCAAGTTAGTAAAGAGTGGCGGCGGATCTT 601

Oy 239 GluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeu 258
Db 600 GAAATAATAAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541

Oy 259 ThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThr 278
Db 540 ACTTATCATAGACGCGCAGCTGGATGATGATGATGATGATGATGATGATGATGATGAT 481

Oy 279 LysHisLeuLeuGluIleValAspLeu 287
Db 480 AACACCTATTAGAAATCGTTGATTTA 454

RESULT 3
AAS54848
ID AAS54848 standard; DNA; 882 BP.
XX
AC AAS54848;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #1160.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR P-PSDB; AAU36989.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX

PS Claim 27; Seq ID No 8485; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 882 BP; 330 A; 116 C; 185 G; 251 T; 0 other;

Alignment Scores:

Pred. No.: 2,32e-189 Length: 882
Score: 197.00 Matches: 268
Percent Similarity: 99.26% Conservative: 0
Best Local Similarity: 99.26% Mismatches: 1
Query Match: 68.64% Indels: 2
DB: 23 Gaps: 0

US-09-925-637-64 (1-287) x AAS54848 (1-882)

QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnGlnLeuSerValAla 20
DB 1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATTAATATCGGTTCG 60

QY 21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
DB 61 ATAATAAATCAGTAATGATCTACTAGCTAGTAGAAGAAGATGTTGTTTCAATTAATGCT 120

QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB 121 GGAGGTAAACGATCCGACGCTCTGTATTACTCCTTATAGATTCACATAATACCGAG 180

QY 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
DB 181 TATGACTTAGGTATGATGAAGCGCAATGCTAGTAATAATGATTCATACATATTCACPTAT 240

QY 81 HisAspAspLeuProAlaMetAspAsnAspTyrArgGlyLysLeuThrAsnHis 100
DB 241 CATGATGACCTACCGATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 300

QY 101 LysValTyrGlyClnThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
DB 301 AAAGTATATGGTGGAGTGGATGCGATATAGCAGGTGCTTATTAACATAAGCATTT 360

QY 121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
DB 361 GAACCTATTTCAAGTGATGATGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 420

QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSer 160
DB 421 CTGTCAATAGCAAGTGTCTGTTGGATGGTGGCGGGTCAATGTTAGATATGCAAGC 480

QY 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
DB 481 GAAGGCCAACCACTTATGCTTGAACCTTTGGAATGATACACAAACAAACAGGAGCA 540

QY 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-Thrly 200
DB 541 CTATTAACTATTCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599

QY 200 sGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspPLe 220
DB 600 AGAACATTTAGAAAGTATAGTTATCATTTAGTATGATGTTTTCAGATTAACATGATTT 659

QY 220 uLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAs 240
DB 660 ATTAGACTGCTATGTTGATGACGAGTATGAGTAAAGTGGCAGGATCTTGAAAA 719

QY 240 nAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTy 260
DB 720 TAATAAAGTACATACGTGAGTTTATAGAAAGATGGCGAGATAAATGACTTA 779

QY 260 rHisArgAspAlaAlaValAspGluLeu 269
DB 780 TCATAGACAGCGCAGCAGTGGATGAACATA 807

RESULT 4
AAS51602
ID AAS51602 standard; DNA; 864 BP.
XX
AC AAS51602;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #19.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR P-PSDB; AAU33743.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 4184; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an

CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 864 BP; 323 A; 112 C; 182 G; 247 T; 0 other;

Alignment Scores:

Pred. No.: 2,56e-184 Length: 864
Score: 192.00 Matches: 263
Percent Similarity: 99.25% Conservativity: 0
Best Local Similarity: 99.25% Mismatches: 1
Query Match: 66.90% Indels: 2
DB: 23 Gaps: 0

US-09-925-637-64 (1-287) x AAS51602 (1-864)

QY 6 MetAsnLysLeuIleAAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
1 ATGAATAAATTAATAGATGAAGTCAATTAATGAATATCGGTGGGATAAATAATACAGTA 60
26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
61 ATGGATCTACAGTACAGAGAAAGTATGTTGATTCATTAATGCTGGAGGTAAACGCATC 120
46 ArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
121 CGACCAAGTCTGTTATTACTACATTTAGATTACATAACCGAGTATGAGTATG 180
66 LysSerAlaIleAlaLeuGluMetIleHisThrYrSerLeuIleHisAspLeuPro 85
181 AAGAGCGCAATTCAGTACAGAAATGATTCATACATATTCATTTATGATGACCTACCA 240
86 AlaMetAspAsnAspTyrArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
241 GCGATGGATATGATGATTCGACGAGGAAATTAACAATCATTAAGTATATGGTGAG 300
106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer 125
301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACCTAAAGCATTTGAATTTCAAGT 360
126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
361 GATGATGATTAATCTGATGAAGTAAATAAATAAAGTTCTACACGCGTGTCAATGCAAGT 420
146 GlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIle 165
421 GGTATGTTGAATGGTGGCGGTCAATGTTAGATATGCAACGCGGAGGCCACCAATTT 480
166 AspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAla 185
481 GATCTTGAACCTTTGGAAATGATACACAAAAACAGGAGCAGCTATTAACTTTTCG 540
186 ValMetSerAlaAlaAspIleAlaAsnValAspThr-ThrLysGluHisLeuGluLeu 205
541 GTTATGATGTCAGCAGATATCGCTAAATGTCGATATGC-AACTAAAGAACATTTAGAAG 599
205 rTyrSerThrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrG1 225
600 TTATAGTATCATTTAGTATGATGATGTTTCAGATTAAAGATGATTTATAGACTGCTATGG 659
225 YAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsnLysSerThrTy 245
660 TGATGAGCGAAGTAGTAAATAAAGTGGCAGCGCATCTTGAAATAATAAAGTACATA 719
245 rValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAl 265
720 CGTCAGTATTATAGGAAAGATGCGCAGAGAGATAAATTTGACTTATCATAGAGCGCAGC 779
265 aValAspGluLeu 269
780 ACTGGATGAACTA 792

RESULT 5

AAS50245/C
ID AAS50245 standard; DNA; 413 BP.

XX
AC AAS50245;

XX 13-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation inhibitory sequence #1469.

XX Antisense; ss: prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 1; Seq ID No 2822; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 413 BP; 115 A; 89 C; 58 G; 151 T; 0 other;

Alignment Scores:

Pred. No.: 4,55e-129 Length: 413
Score: 137.00 Matches: 137
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 47.74% Indels: 0
DB: 23 Gaps: 0

US-09-925-637-64 (1-287) x AAS50245 (1-413)

PS Claim 8; Page 103; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

CC (II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis.

CC (I) and (II) can have antibacterial activity and therefore can be used

CC in vaccination. The nucleic acids (I) may be used to produce the

CC S. epidermidis polypeptides (II) via the production of vectors

CC containing them which are used to produce hosts cells which express the

CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be

CC used to vaccinate subjects and to raise antibodies against the bacteria.

CC The polypeptides may also be used to assay for other inhibitors of their

CC activity and therefore identify compounds that may be used for the

CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to

CC AAH53090 represent specifically claimed S. epidermidis genomic DNA

CC polynucleotide sequences from the present invention. AAH55091 to

CC AAH55098 represent oligonucleotide sequences and primers which are used

CC in the exemplification of the present invention.

CC N.B. The present invention specifically claims all the polynucleotide

CC sequences given in the sequence listing of the present specification,

CC however the sequence listing only goes up to SEQ ID NO:4454 so even

CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,

CC no sequences are present for SEQ ID NO:4455 to 4464.

XX SQ Sequence 783 BP; 297 A; 93 C; 147 G; 246 T; 0 other;

Alignment Scores:

Pred. No.: 8,64e-28 Length: 783

Score: 37.00 Matches: 37

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.89% Indels: 0

DB: 22 Gaps: 0

US-09-925-637-64 (1-287) x AAH52415 (1-783)

QY 70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89

DB 109 GCATTGGAATGATTCATCTATTCTTTAATTCATGATGATTACCAATGGATAAT 168

QY 90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTTP 106

DB 169 GACGATTACCGTAGAGGAAAATTAACAATCATATAAGTTTATGTCGATGG 219

RESULT 9

ABN90871

ID ABN90871 standard; DNA; 909 BP.

XX ABN90871;

24-JUL-2002 (first entry)

Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:334.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

XX antibacterial; gene therapy; gene; ds.

XX Staphylococcus epidermidis.

OS US6380370-B1.

PN 30-APR-2002.

XX 13-AUG-1998; 98US-0134001.

XX 14-AUG-1997; 97US-055779P.

XX 08-NOV-1997; 97US-064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2002-381255/41.

XX P-PSDB; ABP38326.

XX

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis

PT polypeptide, useful for diagnosing and treating bacterial infections -

XX Disclosure; SEQ ID 334; 267pp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading

CC frame (ORF) nucleic acid sequences which encode the amino acid sequences

CC given in ABP35124 to ABP37960. The S. epidermidis sequences have

CC antibacterial activity and can be used in gene therapy. The sequences

CC can also be used in the diagnosis and treatment of bacterial infections,

CC particularly S. epidermidis infections. The sequences can be used to

CC screen for compounds able to interfere with the S. epidermidis life

CC cycle or inhibit S. epidermidis infection.

CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the

CC USPTO web site.

XX SQ Sequence 909 BP; 361 A; 106 C; 165 G; 277 T; 0 other;

Alignment Scores:

Pred. No.: 9,97e-28 Length: 909

Score: 37.00 Matches: 37

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.89% Indels: 0

DB: 24 Gaps: 0

US-09-925-637-64 (1-287) x ABN90871 (1-909)

QY 70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89

DB 235 GCATTGGAATGATTCATCTATTCTTTAATTCATGATGATTACCAATGGATAAT 294

QY 90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTTP 106

DB 295 GACGATTACCGTAGAGGAAAATTAACAATCATATAAGTTTATGTCGATGG 345

RESULT 10

AAH54780

ID AAH54780 standard; DNA; 3444 BP.

XX AAH54780;

03-SEP-2001 (first entry)

S. epidermidis genomic polynucleotide sequence SEQ ID NO:4144.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;

XX vaccination; endocarditis; ds.

XX Staphylococcus epidermidis.

OS WO200134809-A2.

PN 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30782.

XX 09-NOV-1999; 99US-0164258.

XX (GLAX) GLAXO GROUP LTD.

XX Kimmerly WJ;

XX WPI; 2001-316495/33.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,

PT useful for vaccinating against infections, e.g. endocarditis -

XX Claim 8; Page 1844-1845; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

CC

CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC *S. epidermidis* polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed *S. epidermidis* genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX

Sequence 3444 BP; 1365 A; 451 C; 571 G; 1057 T; 0 other;

Alignment Scores:
Pred. No.: 3.58e-27 Length: 3444
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.89% Indels: 0
DB: 22 Gaps: 0

US-09-925-637-64 (1-287) x AAH54780 (1-3444)

Qy 70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAspAsn 89
Db 973 GCATTGGAATGATTCATCTTATTTTAAATTCATGATGATTTACCAAGCAATGGATAAT 1032
Qy 90 AspAspTyrArgGlyGlyLeuThrAsnHisLysValTyrGlyGluTyr 106
Db 1033 GACGATTACCGTAGAGGAAATTAACAATCATAAAGTTTATGTTGTAATGG 1083

RESULT 11

AAH54634/C
ID AAH54634 standard; DNA; 3454 BP.

XX AC AAH54634;

XX DT 03-SEP-2001 (first entry)

XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:3998.

KW *Staphylococcus epidermidis* SRI strain; infection; diagnosis;
XX vaccination; endocarditis; ds.

OS *Staphylococcus epidermidis*.

XX WO200134809-A2.

XX PD 17-MAY-2001.

XX PF 09-NOV-2000; 2000WO-US30782.

XX PR 09-NOV-1999; 99US-0164258.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Kimmerly WJ;

XX DR WPI; 2001-316495/33.

XX Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
XX useful for vaccinating against infections, e.g. endocarditis -

XX Claim 8; Page 1671-1672; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC *S. epidermidis* polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed *S. epidermidis* genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX

Sequence 3454 BP; 1005 A; 574 C; 493 G; 1382 T; 0 other;

Alignment Scores:
Pred. No.: 3.59e-27 Length: 3454
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.89% Indels: 0
DB: 22 Gaps: 0

US-09-925-637-64 (1-287) x AAH54634 (1-3454)

Qy 70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAspAsn 89
Db 150 GCATTGGAATGATTCATCTTATTTTAAATTCATGATGATTTACCAAGCAATGGATAAT 91
Qy 90 AspAspTyrArgGlyGlyLeuThrAsnHisLysValTyrGlyGluTyr 106
Db 90 GACGATTACCGTAGAGGAAATTAACAATCATAAAGTTTATGTTGTAATGG 40

RESULT 12

AAH54565/C

ID AAH54565 standard; DNA; 4045 BP.

XX AC AAH54565;

XX DT 03-SEP-2001 (first entry)

XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:3929.

DE *Staphylococcus epidermidis* SRI strain; infection; diagnosis;
XX vaccination; endocarditis; ds.

OS *Staphylococcus epidermidis*.

XX WO200134809-A2.

XX PD 17-MAY-2001.

XX PF 09-NOV-2000; 2000WO-US30782.

XX PR 09-NOV-1999; 99US-0164258.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Kimmerly WJ;

XX DR WPI; 2001-316495/33.

XX Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
XX useful for vaccinating against infections, e.g. endocarditis -

XX Claim 8; Page 1586-1588; 2188pp; English.
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH53090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX SQ Sequence 4045 BP; 1249 A; 691 C; 536 G; 1569 T; 0 other;

Alignment Scores:
Pred. No.: 4,18e-27 Length: 4045
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.89% Indels: 0
DB: 22 Gaps: 0

US-09-925-637-64 (1-287) x AAH54565 (1-4045)

QY 70 AlaLeuGlutMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAspAsn 89
Db 2106 GCATTGGGAATGATTCATCTATTCTTTAATTCATGATGATTTACGACGAATGGATAAT 2047
QY 90 AspAspTyrArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
Db 2046 GAGGATTACCGTAGAGGAATAATTAACAAATCATATAAGTTATGCGTGAATGG 1996

RESULT 13

AAV15824
ID AAV15824 standard; DNA; 486 BP.

AAV15824;

DT 05-JUN-1998 (first entry)
XX Probe B500 for *Micrococcus luteus* HexPS gene.
DE HexPS gene; hex1; hex2; hex3; prenyl diphosphate synthetase;
XX subunit A; subunit B; polyprenyl diphosphate; vitamin K;
KW ubiquinone; probe; ss.

XX *Micrococcus luteus*.

XX EP812914-A2.

XX 17-DEC-1997.

XX 13-JUN-1997; 97EP-0109692.

XX 14-JUN-1996; 96JP-0154441.

XX (TOYT) TOYOTA JIDOSHA KK.

XX Cho Y, Koike A, Koyama T, Muramatsu M, Ogura K;

XX Shimizu N;

XX

DR WPI; 1998-034975/04.

XX DNA encoding prenyl diphosphate synthetase subunit(s) - new
PT *Micrococcus prenyl diphosphate synthetase* subunit polypeptide(s),
PT and methods for preparing enzymes from subunit(s)

XX Disclosure; Pages 30-31; 46pp; English.

XX The present sequence is a probe for the *Micrococcus luteus* HexPS
CC gene, which comprises three open reading frames designated hex1,
CC hex2 and hex3. hex1 and hex3 are claimed in claims 3 and 3, and
CC encode prenyl diphosphate synthetase (PDS) subunit A and B,
CC respectively.

CC Substances synthesised by PDS, i.e. polyprenyl diphosphates, are
CC precursors of physiologically active substances, e.g. vitamin K and
CC ubiquinones.

XX SQ Sequence 486 BP; 148 A; 85 C; 98 G; 149 T; 5 other;

Alignment Scores:
Pred. No.: 6e-24 Length: 486
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.50% Indels: 0
DB: 19 Gaps: 0

US-09-925-637-64 (1-287) x AAV15824 (1-486)

QY 69 IleAlaLeuGlutMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAsp 88
Db 85 ATTGCGTTGAAATGATTCATGATGATTTAATTCAGGATGACTTGC CGCAATGGAT 144

QY 89 AsnAspTyrArgGlyLysLeuThrAsnHisLys 101
Db 145 AATGATGACTATCGTCGCGGTAAACTCAGCAATCAATAG 183

RESULT 14

ABQ71010

ID ABQ71010 standard; DNA; 5635 BP.

XX AC ABQ71010;

XX 29-AUG-2002 (first entry)

XX *Listeria monocytogenes* 4b contig DNA sequence #952.

XX Antibacterial; *Listeria*; food contamination; mutational analysis;
KW infection; ds.

XX *Listeria monocytogenes* 4b.

XX WO200228891-A2.

XX 11-APR-2002.

XX 04-OCT-2001; 2001WO-FR03061.

XX 04-OCT-2000; 2000FR-0012697.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Kunst F, Glaser P;

XX WPI; 2002-332479/37.

XX New genomic sequences from *Listeria* species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators

XX Claim 14; SEQ ID 3823; 180pp; French.

Query Match: 11.50% Indels: 0
DB: 3 Gaps: 0

US-09-925-637-64 (1-287) x US-09-217-609A-28 (1-486)

QY 69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAsp 88
|||||
DB 85 ATTGGCTTGAATGATTCATGATCTTTTAAATCAGATGACTTCCCGCAATGAT 144
|||||

QY 89 AsnAspAspTyrArgGlyLysLeuThrAsnHisLys 101
|||||
DB 145 AATGATGACTATCGTCGGGTAAACTCAGCAATCATAG 183
|||||

RESULT 4
US-08-873-235B-28
; Sequence 28, Application US/0887235B
; Patent No. 6174715
; GENERAL INFORMATION:
; APPLICANT: MURAMATSU, Masayoshi
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; APPLICANT: SHIMIZU, Naoto
; APPLICANT: CHO, Yenwin
; TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, NW - Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: WordPerfect 6.1 for Windows
; CURRENT APPLICATION DATA:
; FILING DATE: 11-Jun-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: JP 154441/1996
; FILING DATE: 14-Jun-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: TOFFENETTI, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 10235/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-1776
; TELEFAX: 202-429-0796
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 486 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
US-08-873-235B-28

Alignment Scores:
Pred. No.: 2.96e-26 Length: 486
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.50% Indels: 0
DB: 4 Gaps: 0

US-09-925-637-64 (1-287) x US-08-873-235B-28 (1-486)

QY 69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAsp 88
|||||

DB 85 ATTGGCTTGAATGATTCATGATCTTTTAAATCAGATGACTTCCCGCAATGAT 144
|||||

QY 89 AsnAspAspTyrArgGlyLysLeuThrAsnHisLys 101
|||||
DB 145 AATGATGACTATCGTCGGGTAAACTCAGCAATCATAG 183
|||||

RESULT 5
US-08-333-321-1
; Sequence 1, Application US/08333321
; Patent No. 5786192
; GENERAL INFORMATION:
; APPLICANT: Ohta, Shusei
; TITLE OF INVENTION: FARNESYL PYROPHOSPHATE SYNTHETASE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward W. Greason, Esq.
; STREET: 1 Broadway
; CITY: New York
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: No. 5786192epad, Windows 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,321
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/953,424
; FILING DATE: 29-SEP-1992
; APPLICATION NUMBER: JP 3-253788
; FILING DATE: 01-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Greason, Edward W.
; REGISTRATION NUMBER: 18,918
; REFERENCE/DOCKET NUMBER: 077670/00310
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 893 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
US-08-333-321-1

Alignment Scores:
Pred. No.: 9.13e-06 Length: 893
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.88% Indels: 0
DB: 1 Gaps: 0

US-09-925-637-64 (1-287) x US-08-333-321-1 (1-893)

QY 72 GluMetIleHisThrTyrSerLeuIleHisAspLeuPro 85
|||||

DB 226 GAAATGATCCATACGCTACTTTTGATCCATGATGATTGCCG 267
|||||

RESULT 6
US-08-534-910B-2
; Sequence 2, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi

APPLICANT: OBATA, Shusei
APPLICANT: NISHINO, Tokuzo
APPLICANT: OHNUMA, Shinichi
APPLICANT: NAKAZAWA, Takeshi
APPLICANT: OGURA, Kyoza
APPLICANT: KOYAMA, Tanetoshi
TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable
TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Therefor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: U.S.
ZIP: 20036-5405
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.25" Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: IBM/Word Perfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/534,910B
FILING DATE: 28-SEPT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-25253
FILING DATE: 14-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 77670/398
TELEPHONE: (202)429-1776
TELEFAX: (202)429-0796
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
ORIGINAL SOURCE:
ORGANISM: Bacillus stearothermophilus
US-08-534-910B-2

Alignment Scores:
Pred. No.: 9.14e-06 Length: 894
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.88% Indels: 0
DB: 1 Gaps: 0

US-09-925-637-64 (1-287) x US-08-534-910B-2 (1-894)

QY 72 GluMetIleHisThrTyrSerLeuIleHisAspLeuPro 85
|||||
Db 226 GAAATGATCCATGACGTACTTTTGATCCATGATGATTCGCG 267

RESULT 7

US-08-534-910B-3
Sequence 3, Application US/08534910B
Patent No. 5766911

GENERAL INFORMATION:

APPLICANT: KOIKE, Ayumi
APPLICANT: OBATA, Shusei
APPLICANT: NISHINO, Tokuzo
APPLICANT: OHNUMA, Shinichi
APPLICANT: NAKAZAWA, Takeshi
APPLICANT: OGURA, Kyoza
APPLICANT: KOYAMA, Tanetoshi

TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable

TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Therefor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: U.S.
ZIP: 20036-5405
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.25" Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: IBM/Word Perfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/534,910B
FILING DATE: 28-SEPT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-25253
FILING DATE: 14-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 77670/398
TELEPHONE: (202)429-1776
TELEFAX: (202)429-0796
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
ORIGINAL SOURCE:
ORGANISM: Bacillus stearothermophilus
US-08-534-910B-3

Alignment Scores:
Pred. No.: 9.14e-06 Length: 894
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.88% Indels: 0
DB: 1 Gaps: 0

US-09-925-637-64 (1-287) x US-08-534-910B-3 (1-894)

QY 72 GluMetIleHisThrTyrSerLeuIleHisAspLeuPro 85
|||||
Db 226 GAAATGATCCATGACGTACTTTTGATCCATGATGATTCGCG 267

RESULT 8

US-08-534-910B-5
Sequence 5, Application US/08534910B
Patent No. 5766911

GENERAL INFORMATION:

APPLICANT: KOIKE, Ayumi
APPLICANT: OBATA, Shusei
APPLICANT: NISHINO, Tokuzo
APPLICANT: OHNUMA, Shinichi
APPLICANT: NAKAZAWA, Takeshi
APPLICANT: OGURA, Kyoza
APPLICANT: KOYAMA, Tanetoshi

TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable

TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Therefor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, N.W., Suite 600
CITY: Washington
STATE: DC

```
;
; COUNTRY: U.S.
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: IBM/Word Perfect 6.1 Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/534,910B
; FILING DATE: 28-SEPT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: JP 7-25253
; FILING DATE: 14-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Toffenetti, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 77670/398
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)429-1776
; TELEFAX: (202)429-0796
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
; US-08-534-910B-5
;
; Alignment Scores:
; Pred. No.: 9.14e-06 Length: 894
; Score: 14.00 Matches: 14
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 4.88% Indels: 0
; DB: Gaps: 0
;
; US-09-925-637-64 (1-287) x US-08-886-466-1 (1-894)
;
; QY 72 GluMetIleHisThrTyrSerLeuIleHisAspLeuPro 85
; DB 226 GAATGATCCATACGACTCTTTGATCCATGATGATTCGCG 267
;
; RESULT 10
; US-09-475-304-1
; Sequence 1, Application US/09475304
; Patent No. 6225096
; GENERAL INFORMATION:
; APPLICANT: Narita, Keishi
; APPLICANT: Ishida, Chika
; APPLICANT: Takeuchi, Yoshie
; APPLICANT: Ohto, Chikara
; APPLICANT: Ohnuma, Shinichi
; APPLICANT: Nishino, Tokuzo
; TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
; FILE REFERENCE: 77670/494
; CURRENT APPLICATION NUMBER: US/09/475,304
; CURRENT FILING DATE: 1999-12-30
; EARLIER APPLICATION NUMBER: JP 8-191635
; EARLIER FILING DATE: 1996-07-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(891)
; US-09-475-304-1
;
; Alignment Scores:
; Pred. No.: 9.14e-06 Length: 894
; Score: 14.00 Matches: 14
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 4.88% Indels: 0
; DB: Gaps: 0
;
; US-09-925-637-64 (1-287) x US-09-475-304-1 (1-894)
;
; QY 72 GluMetIleHisThrTyrSerLeuIleHisAspLeuPro 85
; DB 226 GAATGATCCATACGACTCTTTGATCCATGATGATTCGCG 267
;
; RESULT 11
; US-09-101-126-4
; Sequence 4, Application US/09101126
; Patent No. 6316216
; GENERAL INFORMATION:
; APPLICANT: OHTO, CHIKARA
; APPLICANT: NAKANE, HIROYUKI
; APPLICANT: NISHINO, TOKUZO
; APPLICANT: OHNUMA, SHINICHI
; APPLICANT: HIROOKA, KAZUTAKE
; TITLE OF INVENTION: MUTATED PRENYL DIPHOSPHATE SYNTHASES
; FILE REFERENCE: 77670/566
; CURRENT APPLICATION NUMBER: US/09/101,126
; CURRENT FILING DATE: 1999-04-27
; EARLIER APPLICATION NUMBER: PCT/JP97/03921
; EARLIER FILING DATE: 1997-10-29
;
; COUNTRY: U.S.
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: IBM/Word Perfect 6.1 Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/534,910B
; FILING DATE: 28-SEPT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: JP 7-25253
; FILING DATE: 14-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Toffenetti, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 77670/398
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)429-1776
; TELEFAX: (202)429-0796
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
; US-08-534-910B-5
;
; Alignment Scores:
; Pred. No.: 9.14e-06 Length: 894
; Score: 14.00 Matches: 14
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 4.88% Indels: 0
; DB: Gaps: 0
;
; US-09-925-637-64 (1-287) x US-08-534-910B-5 (1-894)
;
; QY 72 GluMetIleHisThrTyrSerLeuIleHisAspLeuPro 85
; DB 226 GAATGATCCATACGACTCTTTGATCCATGATGATTCGCG 267
;
; RESULT 9
; US-08-886-466-1
; Sequence 1, Application US/08886466C
; Patent No. 6040165
; GENERAL INFORMATION:
; APPLICANT: Narita, Keishi
; APPLICANT: Ishida, Chika
; APPLICANT: Takeuchi, Yoshie
; APPLICANT: Ohto, Chikara
; APPLICANT: Ohnuma, Shinichi
; APPLICANT: Nishino, Tokuzo
; TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
; FILE REFERENCE: 77670/494
; CURRENT APPLICATION NUMBER: US/08/886,466C
; CURRENT FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: JP 8-191635
; EARLIER FILING DATE: 1996-07-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(891)
; US-08-886-466-1
```

```
; EARLIER APPLICATION NUMBER: JP 8-307506
; EARLIER FILING DATE: 1996-11-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; OTHER INFORMATION: 256-276 is an Asp-rich coding domain
US-09-101-126-4

Alignment Scores:
Pred. No.: 9.14e-06 Length: 894
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.88% Indels: 0
DB: 4 Gaps: 0

US-09-925-637-64 (1-287) x US-09-101-126-4 (1-894)
QY 72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
Db 226 GAAATGATCCATACGACTCTTTGATCCATGATGATTCGCCG 267

RESULT 12
US-09-367-528A-4
; Sequence 4, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
; APPLICANT: TOYOTA JIDOSHA KABUSHIKI KAISHA
; TITLE OF INVENTION: Geranyl Diphosphate Synthetase Gene
; FILE REFERENCE: PH-586
; CURRENT APPLICATION NUMBER: US/09/367,528A
; CURRENT FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: JP97/346686
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(894)
US-09-367-528A-4

Alignment Scores:
Pred. No.: 9.14e-06 Length: 894
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.88% Indels: 0
DB: 4 Gaps: 0

US-09-925-637-64 (1-287) x US-09-367-528A-4 (1-894)
QY 72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
Db 226 GAAATGATCCATACGACTCTTTGATCCATGATGATTCGCCG 267

RESULT 13
US-09-275-742-1
; Sequence 1, Application US/09275742
; Patent No. 6130069
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: ispa
; FILE REFERENCE: GM10205
; CURRENT APPLICATION NUMBER: US/09/275,742
```

```
; CURRENT FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-275-742-1

Alignment Scores:
Pred. No.: 0.014 Length: 876
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: 3 Gaps: 0

US-09-925-637-64 (1-287) x US-09-275-742-1 (1-876)
QY 78 SerLeuIleHisAspAspLeuProAlaMetAsp 88
Db 226 AGCTTGATTCACGATGACCTTCCTGCTATGGAT 258

RESULT 14
US-08-858-207A-173
; Sequence 173, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1284 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-858-207A-173

Alignment Scores:
Pred. No.: 0.0204 Length: 1284
Score: 11.00 Matches: 11
```


Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: 4 Gaps: 0

US-09-925-637-64 (1-287) x US-08-858-207A-173 (1-1284)

QY 78 SerLeulleHisAspLeuProAlaMetasp 88
DB 948 AGCTTGATTCAGGACCTTCCTGCTATGGAT 980

RESULT 15

US-08-961-527-76/c
; Sequence 76, Application US/08961527
; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961.527

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 76:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10011 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-527-76

Alignment Scores:

Pred. No.: 0.156 Length: 10011
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: 4 Gaps: 0

US-09-925-637-64 (1-287) x US-08-961-527-76 (1-10011)

QY 78 SerLeulleHisAspLeuProAlaMetasp 88

DB 7653 AGCTTGATTCAGGACCTTCCTGCTATGGAT 7621

Search completed: May 31, 2003, 00:06:30

Job time : 66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 23:42:12 ; Search time 187 Seconds
(without alignments)
2069.416 Million cell updates/sec

Title: US-09-925-637-64

Perfect score: 287

Sequence: 1 MTNLPNKLIDEVNNELSA.....ELTQIDQFNKHLLEIVDL 287

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

arched: 845702 seqs, 674182571 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1689256

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09925637/runat_23052003.174851_24218/app_query.fasta_1.455
-DB=PublishedApplications.NA -QEXT=fastap -SUFFIX=oli.rnnp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09925637@cgn_1_1_57_erunat_23052003.174851_24218
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications.NA:**

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:**
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:**
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:**
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:**
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:**
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:**
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:**
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:**
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:**
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:**
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:**
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:**
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:**
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	287	100.0	861	9 US-10-084-205-63	Sequence 63, Appl
2	287	100.0	861	10 US-09-925-637-63	Sequence 63, Appl
c 3	287	100.0	1893	7 US-08-781-986A-155	Sequence 155, App
4	197	58.6	882	10 US-09-815-242-8485	Sequence 8485, Ap

5	192	66.9	864	10	US-09-815-242-4184	Sequence 4184, Ap
c 6	137	47.7	413	10	US-09-815-242-2822	Sequence 2822, Ap
c 7	112	39.0	337	10	US-09-815-242-3275	Sequence 3275, Ap
8	18	6.3	882	10	US-09-815-242-6533	Sequence 6533, Ap
9	18	6.3	7528	10	US-09-070-927A-55	Sequence 55, Appl
10	15	5.2	891	9	US-09-941-947A-19	Sequence 19, Appl
11	15	5.2	891	10	US-09-934-903-13	Sequence 13, Appl
12	15	5.2	891	10	US-09-934-868-71	Sequence 71, Appl
13	14	4.9	811	10	US-09-974-300-1015	Sequence 1015, Ap
14	12	4.2	888	10	US-09-815-242-7142	Sequence 7142, Ap
15	12	4.2	900	10	US-09-815-242-5972	Sequence 5972, Ap
16	12	4.2	900	10	US-09-815-242-9986	Sequence 9986, Ap
17	11	3.8	876	10	US-09-815-242-9175	Sequence 9175, Ap
18	11	3.8	876	10	US-09-815-242-9499	Sequence 9499, Ap
19	11	3.8	912	10	US-09-815-242-7285	Sequence 7285, Ap
20	11	3.8	912	10	US-09-815-242-7450	Sequence 7450, Ap
21	10	3.5	474	10	US-09-974-300-5445	Sequence 5445, Ap
22	9	3.1	927	10	US-09-974-300-999	Sequence 999, App
c 23	9	3.1	640681	10	US-09-790-988-1	Sequence 1, Appl
24	8	2.8	235	10	US-09-878-574-13980	Sequence 13980, A
25	8	2.8	436	10	US-09-983-965-176	Sequence 176, App
26	8	2.8	553	10	US-09-070-927A-587	Sequence 587, App
27	8	2.8	586	10	US-09-864-761-12177	Sequence 12177, A
28	8	2.8	623	9	US-10-108-915-19	Sequence 19, Appl
c 29	8	2.8	791	9	US-10-097-065-102	Sequence 102, App
30	8	2.8	888	10	US-09-815-242-7873	Sequence 7873, Ap
31	8	2.8	990	9	US-09-284-320-47	Sequence 47, Appl
32	8	2.8	993	9	US-10-108-915-23	Sequence 23, Appl
33	8	2.8	1062	9	US-10-108-915-15	Sequence 15, Appl
34	8	2.8	1161	9	US-10-108-915-37	Sequence 37, Appl
35	8	2.8	1215	9	US-10-066-500-38	Sequence 38, Appl
36	8	2.8	1215	9	US-10-002-796-38	Sequence 38, Appl
37	8	2.8	1215	9	US-10-066-273-38	Sequence 38, Appl
38	8	2.8	1215	9	US-10-066-494-38	Sequence 38, Appl
39	8	2.8	1215	9	US-10-066-269-38	Sequence 38, Appl
40	8	2.8	1215	9	US-10-066-193-38	Sequence 38, Appl
41	8	2.8	1215	9	US-10-066-211-38	Sequence 38, Appl
42	8	2.8	1224	9	US-09-981-876-22	Sequence 22, Appl
43	8	2.8	1224	9	US-09-148-545-22	Sequence 22, Appl
44	8	2.8	1268	9	US-10-108-915-17	Sequence 17, Appl
45	8	2.8	1316	9	US-09-284-320-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1

US-10-084-205-63
; Sequence 63, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-084-205-63

Alignment Scores:
Pred. No.: 2.02e-291 Length: 861
Score: 287.00 Matches: 287
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-09-925-637-64 (1-287) x US-10-084-205-63 (1-861)

QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
|||||
Db 1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATTAATATATCGGTGG 60
QY 21 IleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAla 40
|||||
Db 61 ATAAATAATCACTAGTACGATGAGTGAAGAAATGATGTTGTTATTAATTAATGCT 120
QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
|||||
Db 121 GGAGGTAACCGATCCGACGATCTGTTATTAATCACTTATGATTAATCAATACCGAG 180
QY 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
|||||
Db 181 TATGAGTTAGGTATGAAGAGCGCAATTCGACTAGAAATGATTCATACATATTCACATT 240
QY 81 HisAspAspLeuProAlaMetAspAsnAspTyrArgGlyLysLeuThrAsnHis 100
|||||
Db 241 CATGATGACCTACCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
|||||
Db 301 AAAGTATATGTCAGTGGATGCGATATAGCAGGTGATGCTTTATTAATAAAGCATTT 360
QY 121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
|||||
Db 361 GAACCTTATTCAGTGATGATGATTAATCACTGATGAAGTAAATAAATTAATTAATTA 420
QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSer 160
|||||
Db 421 CTGTCAATGCAAGTGTGATGTCGAATGTCGCGGTCAATGCTAGATATGCAACG 480
QY 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysAla 180
|||||
Db 481 GAAGGCAACCAATTCGATCTTGAACCTTTGGAATGATACACAAACAAACAGAGCA 540
QY 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrLys 200
|||||
Db 541 TTATTAACCTTTGCGGTTAAGTGCAGCAGATATCGCTAAATGTCGATGATCAACATA 600
QY 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspLeu 220
|||||
Db 601 GAACATTTAGAAAGTTATAGTTATCATTTATGATGATGATGATGATGATGATGATGAT 660
QY 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
|||||
Db 661 TTAGACTGCTATGGTGATGAAGCAAGTTAGGTAAAGTGGGAGCGATCTTGAAGAT 720
QY 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
|||||
Db 721 AATAAAGTACCTAGCTGAGTTATTAAGGAAAGATGCGCCAGAGATTAATTTGACTTAT 780
QY 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
|||||
Db 781 CATAGAGCCAGCAGTGCATCACTAAGCAAAATGATGAACAATTCATCAACAAACAC 840
QY 281 LeuLeuGluIleValAspLeu 287
|||||
Db 841 TTATTAGAAATCGTTGATTTA 861

RESULT 2

US-09-925-637-63
; Sequence 63, Application US/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637

; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-637-63

Alignment Scores:
Pred. No.: 2,02e-291 Length: 861
Score: 287.00 Matches: 287
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-925-637-64 (1-287) x US-09-925-637-63 (1-861)

QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
|||||
Db 1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATTAATATATCGGTGG 60
QY 21 IleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAla 40
|||||
Db 61 ATAAATAATCACTAGTATGATGATCTACGCTAGAGAAAGTATGTTGTTATTAATTAATGCT 120
QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
|||||
Db 121 GGAGGTAACCGATCCGACGATCTGTTATTAATCACTTATGATTAATCAATACCGAG 180
QY 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
|||||
Db 181 TATGAGTTAGGTATGAAGAGCGCAATTCGACTAGAAATGATTCATACATATTCACATT 240
QY 81 HisAspAspLeuProAlaMetAspAsnAspTyrArgGlyLysLeuThrAsnHis 100
|||||
Db 241 CATGATGACCTACCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
|||||
Db 301 AAAGTATATGTCAGTGGATGCGATATAGCAGGTGATGCTTTATTAATAAAGCATTT 360
QY 121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
|||||
Db 361 GAACCTTATTCAGTGATGATGATTAATCACTGATGAAGTAAATAAATTAATTAATTA 420
QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSer 160
|||||
Db 421 CTGTCAATGCAAGTGTGATGTCGAATGTCGCGGTCAATGCTAGATATGCAACG 480
QY 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysAla 180
|||||
Db 481 GAAGGCAACCAATTCGATCTTGAACCTTTGGAATGATACACAAACAAACAGAGCA 540
QY 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrLys 200
|||||
Db 541 TTATTAACCTTTGCGGTTAAGTGCAGCAGATATCGCTAAATGTCGATGATCAACATA 600
QY 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspLeu 220
|||||
Db 601 GAACATTTAGAAAGTTATAGTTATCATTTATGATGATGATGATGATGATGATGATGAT 660

Qy 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
Db 661 TTAGACTGCTATGCTGATGATGAAGCAAGTTAGGTAAAAAAGTGGCGACGATCTTTGAAAAAT 720
Qy 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
Db 721 AATAAAGTACGTACGTGAGTTATTATAGGAAAGATGGCGAGAGATAAATTGACTTAT 780
Qy 261 HisArgAspAlaAlaValAspGluLeuThrGlnLeuAspGluGlnPheAsnThrLysHis 280
Db 781 CATAGACGCGCAGCGAGTGGATGAACAACTAAGCAAAATTGATGAACAAATCAATACAAAAAC 840
Qy 281 LeuLeuGluLeuValAspLeu 287
Db 841 TTATTAGAAATCGTTGATTTA 861

RESULT 3

US-08-781-986A-155/c
; Sequence 155, Application US/08781986A
; Publication No. US20030054436A1

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA: US/08781, 986A

FILING DATE:

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

; NAME: Benson, Bob

; REGISTRATION NUMBER: 30,446

; REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 155:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1893 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-781-986A-155

Alignment Scores:

Pred. No.: 4.37e-291 Length: 1893
Score: 287.00 Matches: 287
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-925-637-64 (1-287) x US-08-781-986A-155 (1-1893)

Qy 1 MetThrAsnLeuProMetAsnLysLeuLeuAspGluValAsnAsnGluLeuSerValala 20

Db 1314 ATGAGAAATCCCGATGATGAATAAATTAATAGATGAATGAATGAATGATGCGG 1255

Qy 21 IleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAla 40

Db 1254 ATAATAATCAGTAATGGATACAGCTAGAGAAAGTATGTTGTTATTCATTAATGCT 1195
Qy 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
Db 1194 GGAGTAAACGACATCCGACCAAGTCTGTTATTACTCACTTATGATTCATAAATACCGAG 1135
Qy 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
Db 1134 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTATACATATTCATTTAT 1075
Qy 81 HisAspAspLeuProAlaMetAspAspAspTyrArgGlyLysLeuThrAsnHis 100
Db 1074 CATGATGACCTACCAGCGATGGATGAATGATGATTCGACGAGGAAAAATTAACAAATCAT 1015
Qy 101 LysValTyrGlyGluTyrThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
Db 1014 AAGTATATGCTGAGTGGAGCTGCATATTACAGGTGATGCTTTATTAACTAAAGCATTT 955
Qy 121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
Db 954 GAACCTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACACGG 895
Qy 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
Db 894 CTGCAATAGCAAGTGGTCATGTTGGAATGGTCGCGGTCAATGTTAGATATGCAAAAGC 835
Qy 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
Db 834 GAAGGCCAACCAATGATCTTGAACCTTTGGCAATGATACACAAAAACAAACAGGACGA 775
Qy 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspThrThrLys 200
Db 774 TTATTAACCTTTTGGGTTATGATGTCAGAGATATCGCTAATGTCGATGATACAACTAAA 715
Qy 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspLeu 220
Db 714 GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAAGATGATTTA 655
Qy 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
Db 654 TTAGACTGCTATGCTGATGAAGCAAGTTAGTAAAAAAGTGGCGACGATCTTCAAAAT 595
Qy 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
Db 594 AATAAAGTACGTACGTGAGTTTATTAGGAAAGATGGCGAGAGATAAATTCATTTAT 535
Qy 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
Db 534 CATAGACGCGCAGCGATGGATGAACCTAACGCAATTTGATGAACAATTCATACAAAAAC 475
Qy 281 LeuLeuGluIleValAspLeu 287
Db 474 TTATTAGAAATCGTTGATTTA 454

RESULT 4

US-09-815-242-8485
; Sequence 8485, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815, 242

; CURRENT FILING DATE: 2001-03-21


```

> CURRENT APPLICATION NUMBER: US/09/815,242
> CURRENT FILING DATE: 2001-03-21
> PRIOR APPLICATION NUMBER: 60/191,078
> PRIOR FILING DATE: 2000-03-21
> PRIOR APPLICATION NUMBER: 60/206,848
> PRIOR FILING DATE: 2000-05-23
> PRIOR APPLICATION NUMBER: 60/207,727
> PRIOR FILING DATE: 2000-05-26
> PRIOR APPLICATION NUMBER: 60/242,578
> PRIOR FILING DATE: 2000-10-23
> PRIOR APPLICATION NUMBER: 60/253,625
> PRIOR FILING DATE: 2000-11-27
> PRIOR APPLICATION NUMBER: 60/257,931
> PRIOR FILING DATE: 2000-12-22
> PRIOR APPLICATION NUMBER: 60/269,308
> PRIOR FILING DATE: 2001-02-16
> NUMBER OF SEQ ID NOS: 14110
> SEQ ID NO: FastSEQ for Windows Version 4.0
> SEQ ID NO 6533
> LENGTH: 882
> TYPE: DNA
> ORGANISM: Enterococcus faecalis
> FEATURE:
> NAME/KEY: CDS
> LOCATION: (1)...(882)
> US-09-815-242-6533

```

Alignment Scores:		
Pred. No.:	1.78e-09	882
Score:	18.00	18
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	6.27%	Mismatches: 0
DB:	10	Indels: 0
		Gaps: 0

US-09-925-637-64 (1-287) x US-09-815-242-6533 (1-882)

Qy	71	LeuGluMetIleHisThrTyrSerLeuIleHisAspSerLeuProAlaMetAsp	88
Db	211	TTAGAGATGATTTCACGTATTCATTCAATCATGATGATTACCAGCAATGGAT	264

US-09-070-927A-55
; Sequence 55, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:

STEVEN BRASII
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009


```
;
;
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-070-927A-55
Alignment Scores:
Pred. No.: 1,44e-08 Length: 7528
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.27% Indels: 0
DB: 10 Gaps: 0
US-09-925-637-64 (1-287) x US-09-070-927A-55 (1-7528)
QY 71 LeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAsp 88
DB 6852 TTAGAGATGATTCATACGATGATTCATTAATTCATGATGATTTACCAAGCAATGCAC 6905
RESULT 10
US-09-941-947A-19
; Sequence 19, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzostowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odom, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL1903 US NA
; CURRENT FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Methylobionas 16a
US-09-941-947A-19
Alignment Scores:
Pred. No.: 2.51e-06 Length: 891
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.23% Indels: 0
DB: 9 Gaps: 0
US-09-925-637-64 (1-287) x US-09-941-947A-19 (1-891)
QY 77 TyrSerLeuIleHisAspLeuProAlaMetAspAsnAspAsp 91
DB 238 TATTCGCTGATTCACGACGATCTGCCGCCATGGACACACGATGAT 282
RESULT 11
US-09-934-903-13
; Sequence 13, Application US/09934903
; Patent No. US20020102690A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, J. Martin
; APPLICANT: Schenzle, Andreas J.
; APPLICANT: No. US20020102690A1ton, Kelley C.
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Rouviere, Pierre
; APPLICANT: Picataggio, Stephen
; APPLICANT: Cheng, Qiong
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
; FILE REFERENCE: CL1646 US NA
; CURRENT APPLICATION NUMBER: US/09/934,903
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: September 1, 2001
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Methylobionas 16a
; FEATURE:
; OTHER INFORMATION: ORF7
US-09-934-903-13
Alignment Scores:
Pred. No.: 2.51e-06 Length: 891
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.23% Indels: 0
DB: 10 Gaps: 0
US-09-925-637-64 (1-287) x US-09-934-903-13 (1-891)
QY 77 TyrSerLeuIleHisAspLeuProAlaMetAspAsnAspAsp 91
DB 238 TATTCGCTGATTCACGACGATCTGCCGCCATGGACACACGATGAT 282
RESULT 12
US-09-934-868-71
; Sequence 71, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M.
; APPLICANT: Schenzle, Andreas J.
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 71
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Methylobionas 16a
; FEATURE:
; OTHER INFORMATION: ORF7 ISPA
US-09-934-868-71
Alignment Scores:
Pred. No.: 2.51e-06 Length: 891
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 5.23% Indels: 0
DB: 10 Gaps: 0
US-09-925-637-64 (1-287) x US-09-934-868-71 (1-891)
Qy 77 TyrSerLeuIleHisAspLeuProAlaMetAspAspAsp 91
|||||
Db 238 TATCGTGATTACGACGATCTGCGGCGCATGGAACGATGAT 282
|||||
RESULT 13
US-09-974-300-1015
; Sequence 1015, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: Expression
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1015
; LENGTH: 811
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1015
Alignment Scores:
Pred. No.: 2 55e-05 Length: 811
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.88% Indels: 0
DB: 10 Gaps: 0
US-09-925-637-64 (1-287) x US-09-974-300-1015 (1-811)
Qy 72 GluMetIleHisThrTyrSerLeuIleHisAspLeuPro 85
|||||
Db 223 GAAATGATCATACGTATTTCATTAAATCCATGACGACCTCCT 264
|||||
RESULT 14
US-09-815-242-7142
; Sequence 7142, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5972
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(900)
US-09-815-242-5972
```

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7142
; LENGTH: 888
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(888)
US-09-815-242-7142
Alignment Scores:
Pred. No.: 0.00348 Length: 888
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.18% Indels: 0
DB: 10 Gaps: 0
US-09-925-637-64 (1-287) x US-09-815-242-7142 (1-888)
Qy 77 TyrSerLeuIleHisAspLeuProAlaMetAsp 88
|||||
Db 238 TATCCTTAATCAGATGATTACCTGCATGAT 273
|||||
RESULT 15
US-09-815-242-5972
; Sequence 5972, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5972
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(900)
US-09-815-242-5972
```

Alignment Scores: Length: 900
 Pred. No.: 0.00353
 Score: 12.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 4.18%
 DB: 10

US-09-925-637-64 (1-287) x US-09-815-242-5972 (1-900)

Qy 77 TvrSerLeuIleHisAspLeuProAlaMetAsp 88
 Db 235 TACTCATTAATTCATGATGATTACCGCGCAATGGAT 270

Search completed: May 31, 2003, 01:02:16
 Job time : 192 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 22:32:52 ; Search time 1558 Seconds
(without alignments)
2983.379 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 287
Sequence: 1 MTNLPMPKLLDEVNNELSA.....ELTQIDRQFNKHLLEIVDL 287

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308016

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-Q=/cgn2.1/USPTO.spool/US09925637/runat_23052003.174850.24121/app.query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=oli.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -SPART=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_ECGN_1_1906_@runat_23052003.174850.24121 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	24	8.4	4484	17	BH770981	BH770981 LLMtag70
2	10	3.5	451	17	AQ176350	AQ176350 HS_3213_A
3	10	3.5	528	17	AQ924302	AQ924302 RPCI-23-2
4	10	3.5	576	13	BI944945	BI944945 sal24d02
C 5	10	3.5	696	17	AG051077	AG051077 Pan trogl
C 6	9	3.1	273	14	F02981	F02981 HSC1IC102 n
7	9	3.1	365	13	BM399846	BM399846 5009-0-62
8	9	3.1	422	17	BH876624	BH876624 hr32c04_b
C 9	9	3.1	474	14	N29629	N29629 yw67f05_s1
C 10	9	3.1	501	17	AQ186598	AQ186598 HS_3079_B
C 11	9	3.1	514	14	BO616272	BO616272 faa99h01
12	9	3.1	516	12	BF702840	BF702840 MI-P-H2-a
13	9	3.1	546	17	BH785617	BH785617 fzm5013f0
14	9	3.1	563	14	BO602325	BO602325 MI-P-HO-a
15	9	3.1	570	12	BF703521	BF703521 MI-P-HI-a
C 16	9	3.1	573	13	BM095848	BM095848 fv27e04.x
17	9	3.1	617	17	BH445833	BH445833 BOHCF25TF
18	9	3.1	662	9	AL641814	AL641814 AL641814
19	9	3.1	666	13	BI717621	BI717621 1031021AO
20	9	3.1	668	13	BI726915	BI726915 1031088EO
21	9	3.1	675	13	BI723806	BI723806 1031067HO
C 22	9	3.1	685	13	BI717620	BI717620 1031021AO
23	9	3.1	685	13	BM449136	BM449136 DSA032E03
24	9	3.1	701	10	BE238154	BE238154 894040E07
25	9	3.1	722	13	BM597369	BM597369 170006874
C 26	9	3.1	783	17	BH706927	BH706927 BOMTK63TF
27	9	3.1	854	17	CNS034T9	CNS034T9 Tetraodon
C 28	9	3.1	855	17	AZ207859	AZ207859 SP_0135_A
C 29	9	3.1	1389	14	BM808194	BM808194 AGENCOURT
C 30	8	2.8	114	17	AZ849898	AZ849898 2M0151101
31	8	2.8	128	17	BH077347	BH077347 RPCI-24-3
C 32	8	2.8	184	17	AZ874342	AZ874342 2M0188L21
33	8	2.8	197	10	BB191918	BB191918 BB191918
34	8	2.8	202	12	BE827852	BE827852 KC2-ET002
35	8	2.8	216	10	BB588852	BB588852 BB588852
36	8	2.8	243	10	AV346846	AV346846 AV346846
37	8	2.8	261	10	AV353343	AV353343 AV353343
38	8	2.8	264	9	AI846208	AI846208 UT-M-AM1-
C 39	8	2.8	266	10	AV343488	AV343488 AV343488
C 40	8	2.8	272	14	F02986	F02986 HSC1IE122 n
C 41	8	2.8	273	17	AZ758116	AZ758116 1M0550A02
42	8	2.8	274	14	T43543	T43543 6806 Lambda
43	8	2.8	289	9	AV146579	AV146579 AV146579
44	8	2.8	297	10	BB192476	BB192476 BB192476
45	8	2.8	300	10	AV325043	AV325043 AV325043

ALIGNMENTS

RESULT 1
BH770981
LOCUS BH770981 484 bp DNA linear GSS 01-MAY-2002
DEFINITION LLMtag706 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, DNA sequence.
ACCESSION BH770981
VERSION BH770981.1 GI:20373938
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 4484)
AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis

JOURNAL
COMMENT

Sci. Alliments, (2002) In press
Contact: Sorokin A
Genetique Microbienne
INRA
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is xsea (94%)
Class: shotgun
High quality sequence start: 30
High quality sequence stop: 4456.
Location/Qualifiers
1. .4484
/organism="Lactococcus lactis subsp. cremoris"
/strain="MG1363"
/db_xref="taxon:1359"
/clone_lib="MG1363 Random Sequence Tag Library"
/notes="Vector: pSGM02; Site_1: SmaI; Library of
chromosomal fragments of L.lactis strain MG1363 was
prepared by partial AluI digestion or by sonication."
BASE COUNT 1540 a 637 c 1023 g 1284 t

FEATURES
source

Alignment Scores:
Pred. No.: 2,33e-11 Length: 4484
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8,36% Indels: 0
DB: 17 Gaps: 0
US-09-925-637-64 (1-287) x BH770981 (1-4484)

Alignment Scores:

Pred. No.: 2,33e-11 Length: 4484
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8,36% Indels: 0
DB: 17 Gaps: 0

US-09-925-637-64 (1-287) x BH770981 (1-4484)

QY 78 SerLeuIleHisAspLeuProAlaMetAspAspAspTyrArgArgGlyLysLeu 97

Db 3867 TCGTTAATCATGACACTGCTCGATGGACATGATGATTCGTGCGAAAGTTA 3926

QY 98 ThrAsnHisLys 101

Db 3927 ACGAACCATATAA 3938

RESULT 2

AQ176350 451 bp DNA linear GSS 17-OCT-1998
LOCUS HS_3213_AL_E01_MR CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate-3213 Col-1 Row-I, DNA sequence.

ACCESSION

AQ176350.1 GI:3573717

VERSION

GSS.

KEYWORDS

human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 451)

REFERENCE

AUTHORS

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

99380589

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3213 row: I column: 1

Class: BAC ends

High quality sequence stop: 451.

FEATURES
source

Location/Qualifiers
1. .451
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-3213 Col-1 Row-I"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT 110 a 102 c 80 g 155 t 4 others
ORIGIN

Alignment Scores:

Pred. No.: 22.7 Length: 451
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.48% Indels: 0
DB: 17 Gaps: 0
US-09-925-637-64 (1-287) x AQ176350 (1-451)

QY 49 LeuLeuLeuThrLeuAspSerLeuAsn 58

Db 381 CTCCTTACTCTTACTCTTCTGATTCACCTGAAT 410

RESULT 3

AQ924302 528 bp DNA linear GSS 21-DEC-1999
LOCUS RPCI-23-278L8.TJ RPCI-23 Mus musculus genomic clone RPCI-23-278L8,
DEFINITION DNA sequence.

ACCESSION AQ924302

VERSION AQ924302.1 GI:6613305

KEYWORDS

GSS.

SOURCE

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 528)

REFERENCE

AUTHORS

Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret

B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.

and Fraser,C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)

or from Resea ch Genetics (info@resgen.com). BAC end page:

http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html

Plate: 278 row: L column: 8

Seq primer: S96

Class: BAC ends.

Location/Qualifiers

1. .528

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-278L8"

/clone_lib="RPCI-23"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:

EcORI; Site_2: EcORI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EcORI and EcoRI Methylase. Size

selected DNA was cloned into the pBac3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). *

BASE COUNT

103 a 146 c 118 g 158 t 3 others

ORIGIN

Alignment Scores: 27.8 Length: 528
 Pred. No.: 10.00 Matches: 10
 Score: 10.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.48% Indels: 0
 DB: 17 Gaps: 0

US-09-925-637-64 (1-287) x AQ924302 (1-528)

Qy 48 ValLeuLeuLeuLeuThrLeuAspSerLeu 57

Db 405 GTATTGCTCTCCGACTTTGGACTCCCTA 434

SULT 4

944945

LOCUS

DEFINITION B1944945 576 bp mRNA linear EST 30-NOV-2001
 sa124802.y1 Gm-cl053 Glycine max cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-cl053-4012 5' similar to TR:Q42698 Q42698 GERANYLGERANYL
 PHOSPHATASE SYNTHASE ; mRNA sequence.

ACCESSION

VERSION B1944945.1 GI:16282306

KEYWORDS

EST.

SOURCE

ORGANISM soybean.

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE

1 (bases 1 to 576)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V., Khanna
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 R., Waterston, R. and Wilson, R.
 Public Soybean EST Project

TITLE

Unpublished (1999)

JOURNAL

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Reggen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 Seq primer: -40RP from Gibco

High quality sequence stop: 391.

FEATURES

source

1. 576
 Location/Qualifiers
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl053-4012"
 /clone_lib="Gm-cl053"
 /tissue_type="Whole seedling, 3 week old, greenhouse
 grown"
 /lab_host="DH10B"

/note="Vector: pBluescript II SKT; Site_1: EcoRI; Site_2:
 XhoI; The Harosoy NIL was constructed and seed was
 provided by Dr. J. Specht, University of Nebraska
 (Shoemaker and Specht, 1995). The cDNA library was
 constructed from mRNA isolated from whole seedlings of 3
 week old greenhouse grown plants. Complementary DNA was
 synthesized from mRNA using a primer consisting of a
 poly(dT) sequence with a XhoI restriction site and a 3'

anchor. EcoRI adapters were ligated to the blunt-ended
 cDNA fragments followed by XhoI digestion. The cDNA
 fragments were directionally cloned into the EcoRI-XhoI
 restriction site of the pBluescript vector. The ligated
 cDNA fragments were transformed into DH10B host cells
 (GibcoBRL). This library was constructed in cooperation
 with Dr. Paul Keim's laboratory at Northern Arizona
 University."

BASE COUNT 112 a 210 c 144 g 110 t

ORIGIN

Alignment Scores: 31.2 Length: 576
 Pred. No.: 10.00 Matches: 10
 Score: 10.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.48% Indels: 0
 DB: 13 Gaps: 0

US-09-925-637-64 (1-287) x B1944945 (1-576)

Qy 40 AlaGlyGlyLysArgIleArgProValLeu 49

Db 265 GCCGGCGAAGAGGATCCGCCCGTCTC 294

RESULT 5

AG051077/c

LOCUS

DEFINITION Pan troglodytes DNA, clone: PTB-031L05.F, genomic survey sequence.

ACCESSION

AG051077

VERSION

AG051077.1 GI:16587969

KEYWORDS

GSS.

SOURCE

Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male

BAC Library clone:PTB-031L05.F.

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.

1

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of Library PTB

Unpublished

2 (bases 1 to 696)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail: chimps@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,

Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end

clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI.

Location/Qualifiers

1. 696

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/clone="PTB-031L05.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT 208 a 133 c 118 g 237 t

ORIGIN

Alignment Scores: 39.8 Length: 696

Pred. No.:

Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.48% Indels: 0
 DB: 17 Gaps: 0

US-09-925-637-64 (1-287) x AG051077 (1-696)

QY 276 PheAsnThrIyHisLeuLeuGluIleVal 285
 Db 201 TTTAACACCAAGCATTTATTGGAATAGTA 172

RESULT 6

F02981/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL
COMMENT

Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@esl.org
Plate: hr32 row: c column: 04
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 422.

FEATURES
source

1..422
/organism="zebra mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="hr32c04"
/lab_host="JM107 or DH5a"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(.x/y reads in M13mp19, .b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."

BASE COUNT 127 a 66 c 82 g 147 t
ORIGIN
Alignment Scores:
Pred. No.: 185 Length: 422
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.14% Indels: 0
DB: 17 Gaps: 0

US-09-925-637-64 (1-287) x BH876624 (1-422)

QY 241 AsnLysSerThrTyrValSerLeuLeu 249
DB 162 AATAGTCGACCTATGTCGTGCTT 188

RESULT 9
N29629/c

LOCUS N29629 474 bp mRNA linear EST 05-JAN-1996
DEFINITION Yv67f05.s1 Soares_placenta_8to9weeks_2NDHP8to9W Homo sapiens cDNA
clone IMAGE:257313 3', mRNA sequence.

ACCESSION N29629
VERSION N29629.1 GI:1148149

KEYWORDS EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project

JOURNAL
COMMENT

Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
High quality sequence stops: 405

Source: IMAGE Consortium; LLNL
This clone is available royalty-free through LLNL; contact the

FEATURES
source

1..474
/organism="Homo sapiens"
/db_xref="GDB:3886923"
/db_xref="taxon:9606"
/clone="IMAGE:257313"
/dev_stage="two placenta; one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: placenta; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(GT) primer [5'
TGTTACCACTCTGAAGTCGAGCGCGCGATTTTTTTTTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 131 a 120 c 119 g 99 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 215 Length: 474
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.14% Indels: 0
DB: 14 Gaps: 0

US-09-925-637-64 (1-287) x N29629 (1-474)

QY 51 LeuLeuThrLeuAspSerLeuAspThr 59
DB 225 TTGCTACGCTTGACTCACTTAACACT 199

RESULT 10
AQ186598/c

LOCUS AQ186598 501 bp DNA linear GSS 01-NOV-1998
DEFINITION HS_3079_B1_B09_T7 CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-3079 Col-17 Row-D, DNA sequence.

ACCESSION AQ186598
VERSION AQ186598.1 GI:3586040

KEYWORDS GSS.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 501)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

JOURNAL 99380589

MEDLINE Contact: Mahairas GG, Wallace JC, Hood L

COMMENT High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence tagged Connector

Plate: 3079 row: D column: 17

Class: BAC ends

High quality sequence stop: 501.

Location/Qualifiers

1..501

/organism="Homo sapiens"

/db_xref="taxon:9606"
 /clone="Plate-3079 Col-17 Row-D"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /sex="male"
 /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
 E-Coli DH10B"
 BASE COUNT 120 a 111 c 93 g 175 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 231 Length: 501
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.14% Indels: 0
 DB: 17 Gaps: 0

US-09-925-637-64 (1-287) x AQ186598 (1-501)

235 GlySerAspLeuGluAsnAsnLysSer 243
 |||||
 420 GCGAGTGATCTCGAAATAATAAGAGC 394

RESULT 11
 BQ616272/c
 LOCUS
 DEFINITION BQ616272 514 bp mRNA linear EST 26-JUN-2002
 faa99h01.x1 Sugano SJD adult male Danio rerio CDNA clone 6034033
 3', mRNA sequence.

ACCESSION BQ616272
 VERSION BQ616272.1 GI:21605941

KEYWORDS EST.
 SOURCE zebrafish.

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 i; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 514)

AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
 S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood
 K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
 Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,
 Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
 and Wilson, R.

TITLE WashU Zebrafish EST Project 1998

JOURNAL

COMMENT

Unpublished (1998)
 Contact: Stephen L. Johnson
 Washington University School
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
 Fax: 314 286 1810

Email: zbrfish@watson.wustl.edu
 Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
 Sequencing by: Washington University Genome Sequencing Center Clone
 Distribution Information can be found through the I.M.A.G.E.
 Consortium/LLNL, send email to: info@image.llnl.gov

zebrafish identity (P-value greater than 1e-99) found to:
 gi215948221gb1AA658668|AA658668 fa56f02.s1 zebrafish gridded kidney
 Danio

Seq primer: T7 from Gibco

High quality sequence stop: 513.

FEATURES
 source

Location/Qualifiers
 1..514

/organism="Danio rerio"

/db_xref="taxon:7955"

/clone="6034033"

/clone_lib="Sugano SJD adult male"

/sex="male"

/tissue_type="whole body"

/dev_stage="adult"

/lab_host="DH10B (phage resistant)"

/note="Vector: pME18S-FL3; Site 1: DraIII (CACCATGTC);

Site 2: DraIII (CACTGTGTC); 1st strand cDNA was primed

with an oligo(dT) primer [ATGTGGCCCTTTTTTTTTTTTTTTT];

double-stranded cDNA was ligated to a DraIII adaptor
 [TGTGGCTACTGG], digested and cloned into distinct DraIII
 sites of the pME18S-FL3 vector (5' site CACTGTGTC, 3' site
 CACCATGTC). XhoI should be used to isolate the cDNA
 insert. Size selection was performed to exclude fragments
 <1.5kb. Library constructed and donated by Dr. Sumio
 Sugano (University of Tokyo Institute of Medical Science).
 Custom primers for sequencing: 5' end primer
 CTCTGCTCTAAAGCTGCG and 3' end primer
 CGACCTGCAGCTCGAGCACA.

BASE COUNT 152 a 98 c 137 g 127 t
 ORIGIN

Alignment Scores:
 Pred. No.: 239 Length: 514
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.14% Indels: 0
 DB: 14 Gaps: 0

US-09-925-637-64 (1-287) x BQ616272 (1-514)

QY 46 ArgProValLeuLeuLeuThrLeu 54

Db 254 AGACGGTCTCTCCTGCTCTTACACTT 228

RESULT 12

BQ702840

LOCUS

DEFINITION

MI-P-H2-abh-d-02-1-UM.s1 MI-P-H2 Sus scrofa cDNA clone

ACCESSION BQ702840

VERSION BQ702840.1 GI:11988248

KEYWORDS EST.

SOURCE pig.

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 516)

AUTHORS

Ronaldo, M.F., Lennon, G. and Soares, M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL

MEDLINE

COMMENT

Contact: Tuggle CK

Molecular Genetics Laboratory, Department of Animal Science

Iowa State University

201 Kildee Hall, Ames, IA 50011-3150, USA

Tel: 5152944252

Fax: 5152942401

Email: cktuggle@iastate.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

non-normalized hypothalamus at estrus day 5 library cDNA library

Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science

Research Center, Department of Animal Science, University of

Missouri-Columbia, 65211 Clone distribution: clones will be

available through Research Genetics (www.resgen.com) The following

repetitive elements were found in this cDNA sequence: 1-57,

>AT rich#Low_complexity 213-248, >AT rich#Low_complexity

Seq primer: M13 Forward

POLYA=yes.

Location/Qualifiers

1..516

/organism="Sus scrofa"

/strain="crossbred"

/db_xref="taxon:9823"

/clone="MI-P-H2-abh-d-02-1-UM"

/clone_lib="MI-P-H2"

/lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not 1; Site_2: EcoRI; The MI-P-H2
 library is derived from hypothalamus at estrus day 5. For
 a detailed description of the library from which this
 clone was derived, please visit our web site at
 http://pigest.genome.iastate.edu/."
 TAG_LIB="MI-P-H2"
 TAG_TISSUE="hypothalamus at estrus day 5"
 TAG_SEQ="TCGTTCG"

BASE COUNT 156 a 82 g 193 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 240 Length: 516
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.14% Indels: 0
 DB: 12 Gaps: 0

US-09-925-637-64 (1-287) x BF702840 (1-516)

QY 198 ThrThrLysGluHisLeuGluSerTyr 206
 DB 449 ACACAAAGACACCTAGATGTCATAC 475

RESULT 13
 BH785617
 LOCUS 546 bp DNA linear GSS 28-NAR-2002
 DEFINITION fzm013f045e05k0 fzm013f045e05 5', DNA sequence.
 ACCESSION fzm013f045e05
 VERSION BH785617
 KEYWORDS GSS.
 SOURCE zea mays.

ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 546)
 AUTHOR Budiman, M.A., Freese, R.G., Bedell, J.A., Nunberg, A.N. and Lakey, N.D.
 TITLE GeneThresher methylation filtered genomic sequences from maize
 JOURNAL Unpublished (2002)
 COMMENT Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: fzm013f045 row: e column: 05
 Seq primer: SK reverse
 Class: shotgun
 High quality sequence stop: 546.

FEATURES
 Source
 1..546
 Location/Qualifiers
 /organism="zea mays"
 /cultivar="MO17"
 /db_xref="taxon:4577"
 /clone="fzm013f045e05"
 /clone_lib="fzm013f045e05"
 /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
 prepared from purified nuclei was randomly sheared,
 end-repaired, size fractionated to enrich for the 0.5 to
 5 kb fraction, ligated into HincII-digested pBCSK(-)
 vector and electroporated into E. coli cells."

BASE COUNT 141 a 102 c 131 g 172 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 258 Length: 546
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.14% Indels: 0
 DB: 17 Gaps: 0

US-09-925-637-64 (1-287) x BH785617 (1-546)

QY 241 AsnLysSerThrTyrValSerLeuLeu 249
 DB 72 AATAAGTCGACCTATGTGTGTTGCTT 98

RESULT 14
 BQ602325
 LOCUS 563 bp mRNA linear EST 24-JUN-2002
 DEFINITION MI-P-HO-ahy-c-10-1-UM.s1 MI-P-HO Sus scrofa cDNA clone
 ACCESSION BQ602325
 VERSION BQ602325.1 GI:21549051
 KEYWORDS EST.
 SOURCE pig.

ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 563)
 AUTHOR Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 9704477
 COMMENT Contact: Tuggle CK

Molecular Genetics Laboratory, Department of Animal Science
 Iowa State University
 201 Kildee Hall, Ames, IA 50011-3150, USA
 Tel: 5152944252
 Fax: 5152942401
 Email: cktuggle@iastate.edu

Tissue Procurement: Dr. Chris Tuggle, Iowa State University
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA
 sequence: 29-59, >AT-rich#low_complexity (matched complement)
 215-251, >AT-rich#low_complexity (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
 Location/Qualifiers
 1..563
 /organism="Sus scrofa"
 /strain="crossbreed"
 /db_xref="taxon:9823"
 /clone="MI-P-HO-ahy-c-10-1-UM"
 /clone_lib="MI-P-HO"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not 1; Site_2: EcoRI; The MI-P-HO
 library is a normalized library derived from hypothalamus
 at estrus days 0 and 12 and ovary at estrus days 0, 5 and
 12. For a detailed description of the library from which
 this clone was derived, please visit our web site at
 http://pigest.genome.iastate.edu/."
 TAG_LIB="MI-P-HO"
 TAG_TISSUE="d_0_hypothalamus"
 TAG_SEQ="TAGATG"

BASE COUNT 177 a 94 c 91 g 201 t
 ORIGIN

Alignment Scores:
 Pred. No.: 269 Length: 563
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.14% Indels: 0

DB: 14 Gaps: 0 DB: 12 Gaps: 0

US-09-925-637-64 (1-287) x BQ602325 (1-563) US-09-925-637-64 (1-287) x BF703521 (1-570)

QY 198 ThrThrLysGluHisLeuGluSerTyr 206 QY 198 ThrThrLysGluHisLeuGluSerTyr 206
||||| 198 ThrThrLysGluHisLeuGluSerTyr 206
Db 452 ACAACAAAGAACACCTAGAGTCATAC 478 Db 453 ACAACAAAGAACACCTAGAGTCATAC 479

RESULT 15

BF703521 570 bp mRNA linear EST 22-DEC-2000
LOCUS MI-P-HI-abm-c-10-1-UM.s1 MI-P-HI Sus scrofa cDNA clone
DEFINITION MI-P-HI-abm-c-10-1-UM 3', mRNA sequence.

ACCESSION BF703521

VERSION

KEYWORDS

SOURCE

plg.

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

Bonaldi.M.F., Lennon.G. and Soares.M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

Contact: Tuggle CK

Molecular Genetics Laboratory, Department of Animal Science

Iowa State University

201 Kildee Hall, Ames, IA 50011-3150, USA

Tel: 5152944252

Fax: 5152942401

Email: cktuggle@iastate.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a Bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

non-normalized hypothalamus at estrus day 0 library cDNA Library

Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science

Research Center, Department of Animal Science, University of

Missouri-Columbia, 65211 Clone distribution: clones will be

available through Research Genetics (www.resgen.com) The following

repetitive elements were found in this cDNA sequence: 1-59,

>AT-rich#Low_complexity 215-252, >AT-rich#Low_complexity

Seq primer: M13 Forward

POLYA=Yes.

FEATURES

Location/Qualifiers

1..570

/organism="Sus scrofa"

/strain="crossbreed"

/db_xref="taxon:9823"

/clone="MI-P-HI-abm-c-10-1-UM"

/clone_lib="MI-P-HI"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-HI

library is derived from hypothalamus at estrus day 0. For

a detailed description of the library from which this

clone was derived, please visit our web site at

http://pigest.genome.iastate.edu/.

TAG_LIB=MI-P-HI

TAG_TISSUE=hypothalamus at estrus day 0

TAG_SEQ=TAGATG*

BASE COUNT 179 a 95 c 93 g 203 t

ORIGIN

Alignment Scores:

Pred. No.: 273 Length: 570

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.14% Indels: 0

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 13:08:18 ; Search time 2483 Seconds
(without alignments)
10091.629 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atacgaatcaccgatgaa.....tattagaatcgttgattta 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_nu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	859.4	99.8	333750	1	AP004827	AP004827 Staphyloc
C 2	854.6	99.3	301550	1	AP003134	AP003134 Staphyloc
C 3	854.6	99.3	346900	1	AP003362	AP003362 Staphyloc
4	853	99.1	882	6	AR106466	AR106466 Sequence
5	449.8	52.2	3444	1	AF270104	AF270104 Staphyloc
6	449.8	52.2	3444	6	AX145422	AX145422 Sequence
C 7	449.8	52.2	4045	1	AF269889	AF269889 Staphyloc
C 8	449.8	52.2	4045	6	AX145207	AX145207 Sequence
9	406.8	47.2	783	6	AX141501	AX141501 Sequence
10	252.6	29.3	250050	1	AL591978	AL591978 Listeria
11	239.8	27.9	5635	6	AX416832	AX416832 Sequence
12	233.4	27.1	1080	1	AB003187	AB003187 Micrococ
C 13	231.2	26.9	195269	6	AX417035	AX417035 Sequence
14	231.2	26.9	33050	1	AL596168	AL596168 Listeria
15	231.2	26.9	349980	6	AX417044	AX417044 Sequence
16	222.6	25.9	2268	1	LMO012349	AX012349 Listeria
C 17	218.6	25.4	309950	1	AP001516	AP001516 Bacillus
C 18	218.4	25.4	11055	1	AE010637	AE010637 Fusobacte
C 19	214.6	24.9	12353	1	AE007710	AE007710 Clostridi
C 20	204	23.7	3454	1	AF269958	AF269958 Staphyloc
C 21	204	23.7	3454	6	AX145276	AX145276 Sequence
C 22	202.4	23.5	296750	1	AP003191	AP003191 Clostridi
C 23	182	21.1	135599	8	CF030821	U30821 Cyanophora
C 24	178.4	20.7	13536	1	AE010066	AE010066 Streptoco
C 25	178.4	20.7	50463	1	AE014159	AE014159 Streptoco
C 26	175.2	20.3	10165	1	AE006583	AE006583 Streptoco
27	174.8	20.3	10029	1	AE006320	AE006320 Lactococc
28	164	19.0	811	6	AX432600	AX432600 Sequence
29	163.2	19.0	348050	1	AP003581	AP003581 Nostoc sp
C 30	161.2	18.7	218470	1	BS000013	Z99116 Bacillus su
31	161.2	18.7	282700	1	BACJH642	D84432 Bacillus su
32	157.2	18.3	894	6	E11869	E11869 Mutated DNA
33	155.8	18.1	1337	6	AX021147	AX021147 Sequence
34	155.6	18.1	894	6	AR148386	AR148386 Sequence
35	155.6	18.1	894	6	E05337	E05337 DNA encodin
36	155.6	18.1	894	6	E11866	E11866 Mutated DNA
37	155.6	18.1	894	6	E11867	E11867 Mutated DNA
38	155.6	18.1	894	6	E11868	E11868 Mutated DNA
39	155.6	18.1	894	6	E14762	E14762 gDNA encodi
40	155.6	18.1	894	6	E16095	E16095 DNA encodin
41	155.6	18.1	894	6	E27521	E27521 Geranyl dip
42	155.6	18.1	894	6	E54864	E54864 Process for
43	155.6	18.1	1260	1	BACFDP5	D13293 B. stearoth
44	154	17.9	894	6	E11865	E11865 Mutated DNA
45	154	17.9	894	6	E27520	E27520 Geranyl dip

ALIGNMENTS

RESULT 1
AP004827/c
LOCUS AP004827 333750 bp DNA linear BCT 02-JUL-2002
DEFINITION Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
strain: MW2, section 6/10.
ACCESSION AP004827 BA000033
VERSION AP004827.1 GI:21204509
KEYWORDS Staphylococcus aureus subsp. aureus MW2 (strain: MW2) DNA.
SOURCE Staphylococcus aureus subsp. aureus MW2
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A.,
Nagai, Y., Iwano, N., Asano, K., Naimi, T., Kuroda, H., Cui, L.,
Yamamoto, K. and Hiramatsu, K.

TITLE Genome and virulence determinants of high virulence community-acquired MRSA
JOURNAL Lancet 359 (9320), 1819-1827 (2002)
MEDLINE 22040717
PUBMED 12044378
REFERENCE 2 (bases 1 to 333750)
AUTHORS Aoki, K., Oguchi, A., Nagai, Y., Asano, K., Iwama, N., Baba, T., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) Akio Oguchi, National Institute of Technology and Evaluation, Biotechnology Center, 2chome 49-10 Nishihara, Shiba-ku, Tokyo 151-0066, Japan
(E-mail: oguchienite.go.jp, URL: http://www.bio.nite.go.jp/, Tel: 81-3-3481-8423, Fax: 81-3-3481-8424)
FEATURES
source Location/Qualifiers
1. .333750
/organism="Staphylococcus aureus subsp. aureus MW2"
/strain="MW2"
/db_xref="taxon:196620"
complement(278. .619)
/gene="MW1341"
complement(278. .619)
/gene="MW1341"
/note="ORFID: MW1341"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB95206.1"
/db_xref="GI:21204510"
/translation="MTNITKEVFDNLEQIEIDVFANKNTLGSSEAKPYLDEYHSKIIDI
FYQNDITGNIDFKLNQPVVPMFKERYMIERYHFMGYROMKFTKTELKMNNA
SQYTRLNKKQV"
complement(624. .1283)
/gene="nth"
complement(624. .1283)
/gene="nth"
/note="ORFID: MW1342"
/codon_start=1
/transl_table=11
/product="endonuclease-like protein"
/protein_id="BAB95207.1"
/db_xref="GI:21204511"
/translation="MYSKALEMIDIVIANMFPDAECLEKHNPELTIIVLLSAQCT
DVLNVRVTEKFTEDYLAVIDEELMNDIRSIGLYRNKAKNIKLCQSLIDQFN
GIPDTHELESAGVKTAVNVMSVAFDEP SLAVDTHVERVSRGLGNRWKDNVRQ
VDRQTSVIPDRWRSHQLIFFGRYHCLARKPKDCICPLEDCREGQKRYKAKE
A"
complement(1273. .1959)
/gene="MW1343"
complement(1273. .1959)
/gene="MW1343"
/note="ORFID: MW1343"
hypothetical protein, similar to chromosome replication
initiation protein dnaD"
/codon_start=1
/transl_table=11
/protein_id="BAB95208.1"
/db_xref="GI:21204512"
/translation="MDKYQLKARPVVRRELLDHYSDGLDRODIAVILLKIYASETS
NKPSIELLGKSTWQPRDITMVIQNLIOLELLOVQKDEGTEYMNLDPFPEKL
SHILKQSWHEKQNSKEKFKOLFVLEDTFARPLSPYSLETNLQWIDVDKHSIIQ
WALDANSLNKLSFYMDRIILLNWKNNKVTIIDSIRKREKFNKPKMTHTVKTVPKFD
WLNGLNLDGK"
complement(2287. .3579)
/gene="asnS"
complement(2287. .3579)
/gene="asnS"
/note="ORFID: MW1344"
/codon_start=1
/transl_table=11
/product="asparaginyl-tRNA synthetase"
/protein_id="BAB95209.1"
/db_xref="GI:21204513"
/translation="MKTITKQAKDHLNDQVTIGAMLTNKRSSGKIAFLQLRDGTGFMQ
GVVKSVEDEEYKLAKEITQESSLYVTGTTINRSDSLGEMQVKSIEVIGSEADHPA
ITPKNGTEFLMDHRLWLRSKQHAVKIRNEVTRATYEFNKQFTVDPDILTAS
APEGSELPHTRYFDQDAFLSOGQLYLEAAMAHCKVFSFGPTPAEKSKYPRHLIE
FWMISGEMAPTNAHSELEIQEQYVHVSVLENCKLELKILERTDTSKLEKATVPPR
ISYDDAIEFLKSGEDDDIEMWEDFGAPHETAANHYDLPVFTINPTSKTKPTWQNP
ENETVLCADLIAPEGYGEIIGSSERVDDLELLEQVKEHGLDEEAYSYYLDRRYGS
VPHCGFGLGERTVAMISGVEHVRETAPPFRLNRLYP"
complement(3901. .6594)
/gene="ding"
complement(3901. .6594)
/gene="ding"
/note="ORFID: MW1345"
/codon_start=1
/transl_table=11
/product="probable ATP-dependent DNA helicase ding"
/protein_id="BAB95210.1"
/db_xref="GI:21204514"
/translation="MGMATYAVVDETTGNQDFFDDIIQIGITFVRNNOIIDYHSMI
RTNLEIPPIQALTSTIEENMLQAPYFNQVAQEIYDKIKDCIFVAHNVDLNFKKA
PKDCNIOVRPKVIDTLEIFKIAFPTDKSYQLSELAEAHGITLANAHRADEAATAK
LMLAFEFKELPDLTKQLYLSKQLYLDIIEFVWROYDAPKLDKSEKFEQII
YRKQVDFKPTNYNGSLASKSAVDQGLYRFQQLYLAETIIDLQMLHSEKAMIEA
SLGSKSLAYLLAALMINIETGKVMISTNTKLLQSKLEKDIIPAMNEALNFKINALL
IKSDSYISGLISQILKDDTSYENILKMLQLLTWITETPSGDSYQLFDDCIVDEA
DQKIETVYPARDVHYNFIRKNAQNIQIGITNHAHLIHSVDENSYQLFDDCIVDEA
HRLPYALNOVTNELSYADIKYQGLIGKNENEKLLKALDQLEKORILEKLDIADID
FCLKASMEIHELNEQLESTIFTIINDSDVDDIHRHNVTFFETKDLKLDLHALID
KLNKLEIFNGISHTKTVKSLRQLLYLKDKNIEQSLKAGHTSFISIKNSQKSTIR
LVVDYAVKDLTKQVLEKFSLIFISGTLKFNHSFEAPKFNKDVHNTFEVNTSL
QSKATSVIFPSDVASYQKNIDEVASIVYIIEYTTTSKCLVFTSYKMMHVVQ
DMNLPEPFEDYVLTQQQNYKIVQFNEDKAILLGTSTFFEGDFQANGIKCVM
IAKLPMKNHAKYWLMDSEFTSTFEYVLPDAVTRFROGLRLIRSENDRCIIVSFD
DRLINSYKNPEQYMLEYRQAKKDIQFGKLLRQIQKKK"
complement(6618. .7589)
/gene="MW1346"
complement(6618. .7589)
/gene="MW1346"
/note="ORFID: MW1346"
/codon_start=1
/transl_table=11
/product="truncated hypothetical protein"
/protein_id="BAB95211.1"
/db_xref="GI:21204515"
/translation="MSKYSDQVLQLLYKNKPNYISGQSTAESLNSRTAVKVIDQLK
LEGKIDSNNHGHLLQOLPDIWYQIIOYTKSSALPDFSEVDSIDSTQAAKSL
VGNQSSFTLSDEQTKGRGNRHHSSKGGGLWMSVLRPNVAESMISKFNLFALG
IRDAIQHSQDEVKVKWPNDIYDNGKVGGLFTEVANNNDGTEALICGIGITQOLE
INDESIRHRATSVQLHDKNKLDRYQFLERLQIEIKRYNQFLTLFSEIRESYIAASN
IWNRTLFTENDKQFQAIDLDYDGYLIVRDEAGESHRLISADIDF"
complement(7576. .8778)
/gene="MW1347"
complement(7576. .8778)
/gene="MW1347"
/note="ORFID: MW1347"
hypothetical protein, similar to poly(A) polymerase"
/codon_start=1
/transl_table=11
/protein_id="BAB95212.1"
/db_xref="GI:21204516"
/translation="MDKSLFEQARPILEQIQDNGFEAYVGVGSVRYVNGRNHIDIDI
TTSAPDDEIRDSIFSTIPVKEHGTINVFNDEYVTTFRAEEDYVDRRPSGVTFV
RDLYEDLQRDRFTMAIMDAYKLYDYFDGQDINNRIIRTVTAERFORFDALRMI
RCLRFOSLSFDIATFEAMRIQADIKFISIRIVITELTKLMRGINVEKSFNHLKS
LKAENYMPYFHELDNQINVTDAIDLELLIAVSVKFDINYSKPLKLSNRQVKDING
YIQIMNALPSITKQKLNQVYDTHLKNVVAADYKNDIOGHEPLIVNLQITID
ETLHLRPLMNRNRMVNGVGLMAHLNKGPMKLDVLRQIEAIVTGVKVSNEETILK
WVDNRHVKI"
complement(8783. .9925)
/gene="MW1348"
complement(8783. .9925)
/gene="MW1348"
/note="ORFID: MW1348"

hypothetical protein, similar to lipopolysaccharide biosynthesis-related pr homolog

```
/codon_start=1
/transl_table=11
/protein_id="BAB95213.1"
/db_xref="GI:21204517"
/translation="MKIGITCYPSMGSSIIATLGLKLAERGHEVHTSNIPRIIR
KPLPNMIFQVEVNYAYFQPPDITLSTKIAIEVKEIGDGLLHMHVAVPHAICGIL
AREMSGDKIKMTLHGDTITVLGYDSLOGAIKFGIEKSDIVSVSLSLAQETHIIF
ETNKEIIPYINFRNEPTKALKSQFGIAPDEKVLIIHVSFROVKRIDTIIETIF
AKREKIPSKLILGDPGLVPMQLTRELNVEDVLFGKQDCVSEFYQLSDVLL
SEKESFGTLLEAKMTGVPIGNSNAGGIKEVKGHGETFVVDGDCDSADYAIRLE
DKVLNKLQKWLADIAERFSGELITDQYEEYVQKMLNEHNSKGE"
complement(10167..10484)
/gene="MW1349"
complement(10167..10484)
/gene="MW1349"
/feature="ORFID:MW1349"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95214.1"
/db_xref="GI:21204518"
/translation="MKSVMVOREVDEYIGQKTYGFSPLANLRLTEVEGELAREIN
HYGEKKKDEEANTIKAEIGDNLFLVLLCLANSMDWTSFNETMEKFFTRDKRF
ERK"
complement(10820..11518)
/gene="MW1350"
complement(10820..11518)
/gene="MW1350"
/feature="ORFID:MW1350"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95215.1"
/db_xref="GI:21204519"
/translation="MSSIFTMIIYFVILMPLWAHQHKVKNYKYSQVSTSGKTGR
EVALDHGANGIYDEVVKGEGFLDHPKRVVLSVSPANDRPSVATIAAHEVG
HAIDQHQHGFYFPAALVPVGNLSSDIYMIIMLIIITLTAIGSFGSTALMIGAGLM
SLAVLFTVTLFVEFDASSRAMKQITALNIVNEKRYKHARKVLSSAAATYVAATAVAV
AELVRIILILRSSD"
```

gene

CDS

gene

CDS

Query Match 99.8%; Score 859.4; DB 1; Length 333750;

Best Local Similarity 99.9%; Pred. No. 4.7e-154;

Matches 860; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGACGAATCTACCGATCAATAAATAATAGATGAAGTCAATAATCAATTCGGTTGGG	60
	118374	ATGAGCATCTACCGATCAATAAATAATAGATGAAGTCAATAATCAATTCGGTTGGG	118315
	61	ATAAATAATCAGTAAATGATCTACGATAGAGAAAGTATGTTCTATTCATTAATGCT	120
Db	118314	ATAAATAATCAGTAAATGATCTACGATAGAGAAAGTATGTTCTATTCATTAATGCT	118255
QY	121	GGAGTAAACGATCCGACCGAGTCTGTTATTCACCTTTAGATTCATTAATACCGAG	180
Db	118254	GGAGTAAACGATCCGACCGAGTCTGTTATTCACCTTTAGATTCATTAATACCGAG	118195
QY	181	TATGATGTTAGTATGAAGCGCAATTCACCTAGAAATGATTCATACATATTCATTT	240
Db	118194	TATGATGTTAGTATGAAGCGCAATTCACCTAGAAATGATTCATACATATTCATTT	118135
QY	241	CATGATGATCTACCGATGATTAATGATGATTAATTCACGAGGAGAAATTAACAATCAT	300
Db	118134	CATGATGATCTACCGATGATTAATGATGATTAATTCACGAGGAGAAATTAACAATCAT	118075
QY	301	AAAGTATATGTTGAGTGCACGCGGATATACGAGGTGATGCTTTTAACTAAACATTT	360
Db	118074	AAAGTATATGTTGAGTGCACGCGGATATACGAGGTGATGCTTTTAACTAAACATTT	118015
QY	361	GAACTTATTTCAAGTGATGATAGATTAATGATGATGAAGTAAATAAATAAATTCACACGG	420
Db	118014	GAACTTATTTCAAGTGATGATAGATTAATGATGATGAAGTAAATAAATAAATTCACACGG	117955

QY	421	CTGTCAATAGCAAGTGGTTCATGTTGAATGGTGGCGGTCAAATGTTAGATATGCAACGC	480
Db	117954	CTGTCAATAGCAAGTGGTTCATGTTGAATGGTGGCGGTCAAATGTTAGATATGCAACGC	117895
QY	481	GAAGCCCAACCAATTCATCTTGAACCTTTGGAAATGATACACAAAACAAACAGGACGA	540
Db	117894	GAAGCCCAACCAATTCATCTTGAACCTTTGGAAATGATACACAAAACAAACAGGACGA	117835
QY	541	TTATTAACCTTTTGGCGTTATAGTGCAGCAGATATCGCTAAATGTCATGATCAACAATAA	600
Db	117834	TTATTAACCTTTTGGCGTTATAGTGCAGCAGATATCGCTAAATGTCATGATCAACAATAA	117775
QY	601	GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTTAAAGATGATTTA	660
Db	117774	GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTTAAAGATGATTTA	117715
QY	661	TTAGACTGCTATGTTGATGAAGCAAAAGTTAGTAAAAAAGTGGCGAGCATCTTGAATAAT	720
Db	117714	TTAGACTGCTATGTTGATGAAGCAAAAGTTAGTAAAAAAGTGGCGAGCATCTTGAATAAT	117655
QY	721	AATAAAGTACCTACCTGAGTTTATTTAGGCAAAAGATGGCGAGAGATAAATGACTTAT	780
Db	117654	AATAAAGTACCTACCTGAGTTTATTTAGGCAAAAGATGGCGAGAGATAAATGACTTAT	117595
QY	781	CATAGAGCGCAGCAGTGGATGAACCTAACCGAAATTTGATGAACAAATTCATCAACAAACAC	840
Db	117594	CATAGAGCGCAGCAGTGGATGAACCTAACCGAAATTTGATGAACAAATTCATCAACAAACAC	117535
QY	841	TTATTAGAAATCGTTGATTTA 861	
Db	117534	TTATTAGAAATCGTTGATTTA 117514	

RESULT 2

AP003134/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

/organism="Staphylococcus aureus subsp. aureus N315"

/strain="N315"

Location/Qualifiers

1. 301550

On Jun 12, 2001 this sequence version replaced gi:13701258.

Tel:81-3-3481-8423, Fax:81-3-3481-8424

(E-mail:oguchi@nate.go.jp, URL:http://www.bio.nite.go.jp/

Nishihara, Shibuva-ku, Tokyo 151-0066, Japan

Technology and Evaluation, Biotechnology Center; 2Chome 49-10

Submitted (30-JAN-2001) Akio Oguchi, National Institute of

Direct Submission

and Kikuchi, H.

Aoki, K., Oguchi, A., Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K.

2. (bases 1 to 301550)

Lancet 357 (9264), 1225-1240 (2001)

Whole genome sequencing of methicillin-resistant Staphylococcus

Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.

Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C.,

Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J.,

Mizutani-Ui, Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C.,

Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,

Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M.,

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,

1


```
CDS
complement(1725..2750)
/gene="SAV1315"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB57477.1"
/db_xref="GI:14247086"
/translation="MISFENDYLEGAHEKVLKRLVDNLVQASGYGFDQFTQAIALEKI
ITIDCPNATIRFLVGGTQTNQVINSMLSEYEGVISADTGHVAVHEGGAIEYSGHKV
ITIPSGKVSADSVETMETKDFKDHVPGMVYISHPTXYGLYSKSELEELC
KVKCYQLPLEMDGARLGLMSQSDMTIKDIAKVCDFYIGGKIGALCEAIYFT
KNERPQFTTRKHGALLAKRLGICQLELFTDNLNISRHAELMANKMKDGFKN
KRLYLFDSPTQQFFILSNIAELAEQVKRFAYWEKYDDQHRVVRFAATSWATTEENL
NKLLELI"
gene
3020..3217
/gene="SAV1316"
CDS
3020..3217
/gene="SAV1316"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57478.1"
/db_xref="GI:14247087"
/translation="MEQIKKFTTAETLELLEKKNINAFUSSEATNKLKLVNITIKIIE
ERTFPNNEEFNAIILTSLVSK"
gene
3238..4755
/gene="SAV1317"
CDS
3238..4755
/gene="SAV1317"
/codon_start=1
/transl_table=11
/product="cardiolipin synthetase homolog"
/protein_id="BAB57479.1"
/db_xref="GI:14247088"
/translation="MEFVPEFFIGKMTYTESNDLGTLETTILAIGFIINLVLAFLII
FLERNRRASSTWANLFLVFLPLGFIILFFGTVSARKLNKNGNVLTFDGLLK
QIESFDGNGYDNKQKHDLVLMDDGDFTENNKVDHFDIDGNDLYDQVLK
IKNAKEYIHELYTFALDGLKRIILHLEPEKLGLEYKLYDDVYSGKNVKNANFDHF
KSLGEVAFATKPLFLNFRNMNHRKIIVIDQLGVGVGNGIDYGLGKLGYW
RDLHRIODGAYDALQLRFLDWSQAHRPOEYDVYFPKKNGLPNSPIQIAASGP
ASDHIQIEGYTKMTSAKKSIVYLOSPIFDPNSVINAIKAAKSGVDVHMLPKCPD
HPLVWAFESNADLLSGVKIITYENGFIHSMKCLIDDEIVSVGTANWDFRSPENF
EVNAYVDENLAKURVAYEHDTITASKQUTKESYANRPLSVKFKESLAKLVSPIL"
4939..5436
/gene="SAV1318"
CDS
4939..5436
/gene="SAV1318"
/note="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57480.1"
/db_xref="GI:14247089"
/translation="MIQISNKNKRCVLSNISFDIEQKCIALIGKNGAGKSTLI
DLIGNWNAGSEIFDKDLQSENRSIMFQKTPDQLKVIETIINLYOSFENPLPL
EEIIEETKSQLNQFNKLGGQQRLLDFVLSLIGQPQLILDEPTSTMDIETREY
FWSII"
gene
5461..5838
CDS
5461..5838
/gene="SAV1319"
/note="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57481.1"
/db_xref="GI:14247090"
/translation="MLYTSHYIEEVRMSDKIILENGEILINDSTSHRTNQSQOIT
LSDEYIRKLKDLKDLVQKNHNGTIKIITSNVNDTILYLOQLHINLDDIETQKVSIV
DSYFNKKQKQSGSNYDKLEENRI"
gene
5807..6538
CDS
5807..6538
/gene="SAV1320"
/note="similar to ABC transporter integral membrane protein"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57482.1"
/db_xref="GI:14247091"
/translation="MILSYLKIEFKVIMRKKTLLILSPFVIFYILFTSILELPEDV
KPFYKEYMYSMTVYSLLSFSLTLPDLIINEKQEMRQRLMTPTFTSYVYSKVVK
TMLQRAIALIVFMVGHGFKGVMSAVQWLESIGIFLWLGASLLIFEGILFSLNDLQK
TALANIVTIGLAVLGGMLPINTPTFNLQHVAVLPSYHLKRLGVDIASNHHNLIIS
FAIILYLVGISTIAVYICSHFKRAE"
6542..7633
/gene="SAV1321"
CDS
6542..7633
/gene="SAV1321"
/note="similar to two-component sensor histidine kinase"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57483.1"
/db_xref="GI:14247092"
/translation="MKFLKDTSAIEISSILYLIFFPIAGIFFNEVGPWKLYIISVIVF
SLSYLILVNNRNLMTFYLIIHIFICYFVSHPMLSLFFYSAPVFPFKN
NVKKTATNLFIITMIICTIITLYLNNYFAMMYVYVLSLMLDNFKMKKREYQKE
IAEKNRHINTLAEORHRIGODLHDTLGHVPSLSKSELAVKLIDTDEVKAEELL
AINKLSRESLNKREIIDDVKLPSRIEEDSRKVLKDADIDFTENKELAQVLSPTK
QSNLYMTRERAINNVIKHANSKGVHKLATVNNHKLMLIEDDGKIDGSDCEVKSISQ
RVQHLNGTLAVDSTNGTKIILIEISTGGIA"
7630..8232
/gene="SAV1322"
CDS
7630..8232
/gene="SAV1322"
/note="similar to two-component response regulator"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57484.1"
/db_xref="GI:14247093"
/translation="WTSLLIAEDQNLQAMVQLIKHGDFFILADTNGLDAMKILIE
ETNPANVLDIEMPGMTGLEIAERKKHLNKIIVITFKRPGYFEKAVVNDVDAVY
LKERSIEELVETINKVNGEKEYSATLMTSFVDKNPLTPKEQIVLEIGNGLSKEI
SEKFLTDGTGTVRNYTVIIDKLFADNRFDAWKANKEKGI"
complement(8345..8533)
/gene="SAV1323"
CDS
complement(8345..8533)
/gene="SAV1323"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="MNISENVTIIVAIILIIIGFIQIFPKPLPIVTAVALSILLGI
FVGFIYIVISFYNNKRN"
8672..9205
/gene="nuc"
CDS
8672..9205
/gene="nuc"
/note="SAV1324"
/codon_start=1
/transl_table=11
/product="thermonuclease"
/protein_id="BAB57486.1"
/db_xref="GI:14247095"
/translation="MKSNRSLAMIVAIIVGVLAFOFMNHTGPPFKGTNHTVODLN
Query Match 99.3%; Score 854.6; DB 1; Length 346900;
Best Local Similarity 99.5%; Pred. No. 3.8e-153;
Matches 857; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGACCAATCTACCGCATGAATAATTAATAGATGAAGTCAATAATATCGGTGCG 60
|||||
```

Db 251159 ATGACGAATCTACCGATGAATAAATAATAGATGAAGTCAATAATGAATATATCGTTGG 251100
QY 61 ATAAATAATCAATAGTAAATGATCTAGCTAGAGAAAGTATGTTGATATCAATTAATGCT 120
Db 251099 ATAAATAATCAATAGTAAATGATCTAGCTAGAGAAAGTATGTTGATATCAATTAATGCT 251040
QY 121 GGAGTAAACGATCCGACCGAGTCTGTTATTAATCACTTATGATTCATTAATACCGAG 180
Db 251039 GGAGTAAACGATCCGACCGAGTCTGTTATTAATCACTTATGATTCATTAATACCGAG 250980
QY 181 TATGAGTTAGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCATCTATT 240
Db 250979 TATGAGTTAGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCATCTATT 250920
QY 241 CATGATGACCTACCGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 250919 CATGATGACCTACCGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 250860
QY 301 AAAGTATATGTTGAGTGGACGCTGATATAGCAGGTGATGCTTTATTAATCACTTAACGATTT 360
Db 250859 AAAGTATATGTTGAGTGGACGCTGATATAGCAGGTGATGCTTTATTAATCACTTAACGATTT 250800
QY 361 GAACCTATTTCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 250799 GAACCTATTTCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250740
QY 421 CTGCAATATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 250739 CTGCAATATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 250680
QY 481 GAAGCCCAACCAATGATCTTGAACCTTTGGAATGATACACAAACAAACAGGAGCA 540
Db 250679 GAAGCCCAACCAATGATCTTGAACCTTTGGAATGATACACAAACAAACAGGAGCA 250620
QY 541 TTATTAACCTTTGCGGTTATGAGTGCAGCATATCGCTAATGTCGATGATCACTATAA 600
Db 250619 TTATTAACCTTTGCGGTTATGAGTGCAGCATATCGCTAATGTCGATGATCACTATAA 250560
QY 601 GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 250559 GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250500
QY 661 TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 250499 TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250440
QY 721 AATAAAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 780
Db 250439 AATAAAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 250380
QY 781 CATAGAGCGCAGCAGTGGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 250379 CATAGAGCGCAGCAGTGGATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 250320
QY 841 TTATTAGAAATCGTTGATTTA 861
Db 250319 TTATTAGAAATCGTTGATTTA 250299

RESULT 4
AR106466
LOCUS AR106466 882 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6107058.
ACCESSION AR106466
VERSION AR106466.1 GI:12820996
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 882)
AUTHORS Gwynn,M. and Wilding,E.Imogen.
TITLE IsPa from Staphylococcus aureus
JOURNAL Patent: US 6107058-A 1 22-AUG-2000;

FEATURES
source
BASE COUNT 328 a 118 c 186 g 250 t
ORIGIN
Query Match 99.1%; Score 853; DB 6; Length 882;
Best Local Similarity 99.4%; Pred. No. 2.2e-152;
Matches 856; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGACGAATCTACCGATGAATAAATAATAGATGAAGTCAATAATGAATATATCGTTGG 60
Db 1 ATGACGAATCTACCGATGAATAAATAATAGATGAAGTCAATAATGAATATATCGTTGG 60
QY 61 ATAAATAATCAATAGTAAATGATCTAGCTAGAGAAAGTATGTTGATATCAATTAATGCT 120
Db 61 ATAAATAATCAATAGTAAATGATCTAGCTAGAGAAAGTATGTTGATATCAATTAATGCT 120
QY 121 GGAGTAAACGATCCGACCGAGTCTGTTATTAATCACTTATGATTCATTAATACCGAG 180
Db 121 GGAGTAAACGATCCGACCGAGTCTGTTATTAATCACTTATGATTCATTAATACCGAG 180
QY 181 TATGAGTTAGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCATCTATT 240
Db 181 TATGAGTTAGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCATCTATT 240
QY 241 CATGATGACCTACCGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 241 CATGATGACCTACCGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 AAAGTATATGTTGAGTGGACGCTGATATAGCAGGTGATGCTTTATTAATCACTTAACGATTT 360
Db 301 AAAGTATATGTTGAGTGGACGCTGATATAGCAGGTGATGCTTTATTAATCACTTAACGATTT 360
QY 361 GAACCTATTTCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 GAACCTATTTCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 CTGCAATATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 421 CTGCAATATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GAAGCCCAACCAATGATCTTGAACCTTTGGAATGATACACAAACAAACAGGAGCA 540
Db 481 GAAGCCCAACCAATGATCTTGAACCTTTGGAATGATACACAAACAAACAGGAGCA 540
QY 541 TTATTAACCTTTGCGGTTATGAGTGCAGCATATCGCTAATGTCGATGATCACTATAA 600
Db 541 TTATTAACCTTTGCGGTTATGAGTGCAGCATATCGCTAATGTCGATGATCACTATAA 600
QY 601 GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 661 TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 AATAAAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 780
Db 721 AATAAAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTAT 780
QY 781 CATAGAGCGCAGCAGTGGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 CATAGAGCGCAGCAGTGGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TTATTAGAAATCGTTGATTTA 861
Db 841 TTATTAGAAATCGTTGATTTA 861
RESULT 5
AP270104

LOCUS	AF270104	3444 bp	DNA	linear	BCT 01-AUG-2000
DEFINITION	Staphylococcus epidermidis strain SRI clone step.1049b08 genomic sequence.				
ACCESSION	AF270104				
VERSION	AF270104.1	GI:9624010			
KEYWORDS					
SOURCE	Staphylococcus epidermidis.				
ORGANISM	Staphylococcus epidermidis				
REFERENCE	Bacteria; Firmicutes; Bacilliales; Staphylococcus.				
AUTHORS	1 (bases 1 to 3444) Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M., Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I., Listenbee,S., Ashanti,C., Altschuller,G., Mamo,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and Furdon,P.J.				
TITLE	transposon-mediated sequencing of the staphylococcus epidermidis genome				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 3444)				
AUTHORS	Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M., Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I., Listenbee,S., Ashanti,C., Altschuller,G., Mamo,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and Furdon,P.J.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA				
FEATURES	Location/Qualifiers				
source	1..3444 /organism="staphylococcus epidermidis" /strain="SRI" /db_xref="taxon:1282" /clone_ref="step.1049b08"				
BASE COUNT	1365 a	451 c	571 g	1057 t	
ORIGIN					
Query Match	52.2%; Score 449.8; DB 1; Length 3444;				
Best Local Similarity	70.2%; Pred. No. 6.4e-76;				
Matches	604; Conservative	0; Mismatches	257; Indels	0; Gaps	0;
QY	1 ATGAGCAATCTCCGGATGAATAAATTAAATAGATGAAGTCAATTAAGAATTATTCGGTTGGG 60				
Db	766 ATGAGAAACTACAGATGAATAAATTAAATATATAATTAATCATCCTGAATAGTCA 825				
QY	61 ATAAATAAATCAGTAAATGGATCTCAGCTAGAGAAGATGTCTGTATTCAAATGCCT 120				
Db	826 ATACAATCATCACCAATTAATAAACCAATTTAGAAGAAGATGCAATATATCAATTAATGCT 885				
QY	121 GGAGGTAACCGATCCGCCAGCTTCGTATTACTCACTTTAGATTCACATAATACCAG 180				
Db	886 GTGGTTAAAGAATCAGACCAGCTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 945				
QY	181 TATGAGTTAGGTATGAAGAGCCCAATTGCACTAGAAAATGATTTCATACATATCACTTATT 240				
Db	946 TATCAACAGGACTAAATAGTCTTTAGCATTTGGAATGATTTCATACITATTCTTTAAT 1005				
QY	241 CATGATGACCTACCGAGCATGGAATGATGATTATCGACGAGGAAAAATTAACAAATCAT 300				
Db	1006 CATGATGATTTACCGACGAATGGAATGACGATTACCGGTAGAGGAAAAATTAACAAATCAT 1065				
QY	301 AAGGTATATGGTGAGTGGACTCCGATATTAGCAGCTGATGCTTTATTAACTAAAGCATTT 360				
Db	1066 AAGGTTTATGGTGAATGGAAAGCCCAATCTTGCTGGTGATGCAATTATTAAACAAAGCTTTT 1125				
QY	361 GAACCTATTTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTTACAACGG 420				
Db	1126 GAATTAGTTTCTTAATGATCACTACCATTGGAAGTAGTGTGAAGTAGTATTATATAAAGA 1185				
QY	421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTGGCGGTCAAAATGTTAGATATGCAAGC 480				
Db	1186 CTTTCAAAGCAAGTGGACATTTGGGAATGGTGGTGGCCAGCGCTGTGATATGGAAGT 1245				

Db 1006 CATGATGATTTACAGCAATGATATGACGATTTACCGTAGAGAAATTAACAATCAT 1065
QY 301 AAAGTATATGTTGAGTGGACTCGGATATAGCAGGTGATGCTTTTAACTAAAGCATTT 360
Db 1066 AAAGTTATGTTGTAATGGAAGCCATTTCTGCTGCTGATGATTAATTAACAAGCTTTT 1125
QY 361 GAACATATTTCAAGTCATGATAGATTAATGATGAAGTAATAAATTAAGATTTCAACAGG 420
Db 1126 GAATGATGTTTAAATGATATACATTTAGAGTAGTGTGAAGTAGTAAATTAATAAAGA 1185
QY 421 CTGTCAATAGCAAGTGGTCAATGTTGAATGTCGCGGTCAAAATGTTAGATATCAAAAGC 480
Db 1186 CTTTCAAAAGCAAGTGGACATTTGGGAATGTTGGGTGGCCAGCGCTTGATATGAAGT 1245
QY 481 GAAGGCCAACCAATGATGTTGAAACTTTTGGAAATGATACACAAAACAAAACAGGAGCA 540
Db 1246 GAAGGGAAGTCAATTCGTTTGAAGAACTTTAGAAATCAATTCATGAAACTTAAGACAGCGCT 1305
QY 541 TTATTAATTTTTCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 1306 TTACTAAATTTTTCAGTTATGCTGCGGTAGACATTTGCTCAAGTAGAACAATAATTTGCT 1365
QY 601 GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 1366 AGAATTTAGATGAATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425
QY 661 TTGACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 1426 CTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1485
QY 721 AATAAAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 1486 CATAAAGTACTTATGTTCTTCTTACTTGGAAAGAGGAGCAGAAAGTAAACAT 1545
QY 781 CATAGAGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 1546 CATCAATATCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
QY 841 TTATTAGAAATCGTTGATTTA 861
Db 1606 TTAAGTGATATGTAGATTTA 1636

RESULT 7
AF269889/c
LOCUS
DEFINITION
Staphylococcus epidermidis strain SRI clone step. 1028f08 genomic
sequences.
AF269889
AF269889.1 GI:9623789
Staphylococcus epidermidis.
Staphylococcus epidermidis
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
AUTHORS
Kimmerly, W.J., Taylor, J., David, N., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelsen, F.J., Rivers, P.R., Torruella-Miller, I.,
Listenbee, S., Ashanti, C., Altshuler, G., Mano, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE
Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL
REFERENCE
AUTHORS
Taylor, J., David, N., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelsen, F.J., Rivers, P.R., Torruella-Miller, I.,
Listenbee, S., Ashanti, C., Altshuler, G., Mano, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
TITLE
Direct Submission
JOURNAL
Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore

FEATURES
source
Drive, Research Triangle Park, North Carolina 27709-3398, USA
1. 4045
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1028f08"
BASE COUNT 1249 a 691 c 536 g 1569 t
ORIGIN
Query Match 52.2%; Score 449.8; DB 1; Length 4045;
Best Local Similarity 70.2%; Pred. No. 6.2e-76;
Matches 604; Conservative 0; Mismatches 257; Indels 0; Gaps 0;
QY 1 ATGACGAATCTACCGATGAATTAATTAATAGATGAAGTCAATTAATGAATATCGTGTGG 60
Db 2313 ATGAAGAAATACAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2254
QY 61 ATAATTAATCACTAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Db 2253 ATACAATCATCACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2194
QY 121 GGAGGTAAACGATCGACCCAGTTCGTGTTTACTTACTCCTTTAGATTCACTTAATACCGAG 180
Db 2193 GGTGTTAAAGAAATCAAGCCAGTCAATTAATTAATTAATTAATTAATTAATTAATTA 2134
QY 181 TATGATTTAGTATGAAGAGCGCAATTCACATAGAAATGATTCATACATATTCACATTT 240
Db 2133 TATCAACAGGACTAAATAGTCTTTAGCATTTGGAATGATTCATCTATTTCTTAAT 2074
QY 241 CATGATGCTTACCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 2073 CATGATGATTTACCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2014
QY 301 AAAGTATATGTTGAGTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 2013 AAAGTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1954
QY 361 GAACATTTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 1953 GAATGATTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1894
QY 421 CTGTCAATAGCAAGTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 1893 CTTTCAAAAGCAAGTGGACATTTGGGAATGTTGGGTGGCCAGCGCTTGATATGAAAGT 1834
QY 481 GAAGGCCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 1833 GAAGGGAAGTCAATTCGTTTGAAGAACTTTAGAAATCAATTCATGAAACTTAAGACAGCGCT 1774
QY 541 TTATTAACTTTTTCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 1773 TTACTAAATTTTTCAGTTATGCTGCGGTAGACATTTGCTCAAGTAGAACAATAATTTGCT 1714
QY 601 GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 1713 AAGAAATTTAGATGAATTTAGTCACTATTTAGGAATGATGTTTCAATTAAGATGATTTA 1654
QY 661 TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 1653 CTGGATGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594
QY 721 AATAAAGTACCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 1593 CATAAAGTACTTATGTTCTTCTTACTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1534
QY 781 CATAGAGCGCAGCAGTGGATGAATCAACGCAATTAATGATGAACAAATTAATTAACAC 840
Db 1533 CATCAATATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1474
QY 841 TTATTAGAAATCGTTGATTTA 861
Db 1473 TTAAGTGATATGTAGATTTA 1453


```

421 CATGAAGTAAGACAGCGCTTACTAAATTTTCAGTATGCGGTACAGATTGCT 480
580 AATGTCGATGATACAACTAAAGAACATTAGAAAAGTTATAGTTATCATTTAGGTATGATG 639
481 CAAGTAGAACAAATATTGCTAAGATTAGATGAATTTAGTCAATTTAGGATGATG 540
640 TTCAGATTAAAGATGATTTATAGACTGCTATGATGATGATGAAGCAAGTTAGTAAAAAA 699
541 TTTCAATTAAAGATGATTTACTGATGATGATGATGATGATGATCAAAACTTGGCAAAAA 600
700 GTGGCAGCGATCTTGAATAATAAAGTACGTACGTGACGTTTATTAGGGAAGATGCG 759
601 GTAGCGATGATATAGTAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 660
760 GCAGAGATTAATGACCTTATCATAGACGACGAGTGATGATGATGATGATGATGATGAT 819
661 GCAGAGAAAGCTTAAACATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 720
820 GAACAATTCAATACAAAACACTTATAGAAATCGTTGATTTA 861
721 GATCAATATGATACCTCTCGAATTAAGTATGATATGATGATTTA 762

RESULT 10
AL591978      250050 bp   DNA   linear   BCT 06-JUN-2002
Listeria monocytogenes strain EGD, complete genome, segment 6/12.
AL591978 AL591824
AL591978.1 GI:16410540
Listeria monocytogenes.
Listeria monocytogenes
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Anand, A.,
Baquero, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T.,
Charbit, A., Chetoui, F., Couve, E., de Daruvar, A., Dehouk, P.,
Dommann, E., Dominguez-Bernal, G., Duchaud, E., Durand, L.,
Dussurget, O., Entian, K. D., Fsihi, H., Portillo, F. G., Garrido, P.,
Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J.,
Jackson, D., Jones, L. M., Kaerst, U., Kref, J., Kuhn, M., Kunst, F.,
Kurapat, G., Madieno, E., Maitournam, A., Vicente, J. M., Ng, E.,
Nedjari, H., Nordstieck, G., Novella, S., de Pablo, B., Perez-Diaz, J. C.,
Purcell, R., Remmel, B., Rose, M., Schluter, T., Simoes, N.,
Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. and Cossart, P.
Comparative genomics of Listeria species
Science 294 (5543), 849-852 (2001)
2157279
1167969
2 (bases 1 to 250050)
Glaser, P., Frangeul, L. and Rusniok, C.
Direct Submission
Submitted (06-JUN-2001) Glaser, P., Institut Pasteur, Genomique des
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cedex 15, FRANCE
E-mail: pglaser@pasteur.fr
Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
Location/Qualifiers
1. 250050
/organism="Listeria monocytogenes"
/strain="EGD-e"
/db_xref="taxon:1639"
210..753
/gene="lmo1125"
210..215
/gene="lmo1125"
220..753
/codon_start=1
/protein_id="CAC99203.1"
/db_xref="GI:16410541"

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
E-mail: pglaser@pasteur.fr
Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
Location/Qualifiers
1. 250050
/organism="Listeria monocytogenes"
/strain="EGD-e"
/db_xref="taxon:1639"
210..753
/gene="lmo1125"
210..215
/gene="lmo1125"
220..753
/codon_start=1
/protein_id="CAC99203.1"
/db_xref="GI:16410541"

FEATURES
source
gene
RBS
CDS

```

```

/db_xref="SPTREMBL:O8Y724"
/translation="MLPSYLSKSTAKELSIINDFLNVEIVTINKEETFEILYCGTLKEI
EGDLITGEAYEIPKIIIAKSLTSGKEILLYDGYGYDSMECFDEFEERATQNRLEKK
YPINLSNRLSIGIGIDYSEKEEVEFDENGHVLLIDDRHPWQVKTDGDFLEIT
ATDNGDSLLILTEELA"
join(764..769,778..1230)
/gene="lmo1126"
764..769
/gene="lmo1126"
778..1230
/gene="lmo1126"
/note="similar to E. coli YjAB protein"
/codon_start=1
/transl_table=11
/protein_id="CAC99204.1"
/db_xref="GI:16410542"
/db_xref="SPTREMBL:O8Y723"
/translation="MILINKRAENQDYQMLAIWEKSVIATHDFLAAEDRQYKQIIP
LFDNVEILFWPAGIEIVGFSCTSERELDMFLNPATNGYGSQILLIWLLENKRLNI
VDVNPQENATRFYLGKHFVSVSSRDLGDFGNPYFILHLQKERVNGI"
1208..1213
/gene="lmo1127"
1208..1213
/gene="lmo1127"
1220..1692
/gene="lmo1127"
1220..1675
/gene="lmo1127"
/codon_start=1
/transl_table=11
/protein_id="CAC99205.1"
/db_xref="GI:16410543"
/db_xref="SPTREMBL:O8Y722"
/translation="MEYELQLONELELEIIOEQLQNEKRGIMVKTAKDGTITRL
AKTKDKAGYFVAFWEKDEFGCEPDENDCADFLCIVVDGLNGLFVPEQCLIDK
GILTSEASIGKMAARPYTWCQDLNKTAKTQAOQVYFKNYSIKREDH"
1660..1692
/gene="lmo1127"
/complement(1689..2711)
/gene="lmo1128"
/complement(1689..2711)
/gene="lmo1128"
/codon_start=1
/transl_table=11
/protein_id="CAC99206.1"
/db_xref="GI:16410544"
/db_xref="SPTREMBL:O8Y721"
/translation="MAKNGVDIIRKKKTIIMSALIIVLVAIFIGNYEMNETRVA
IPTGKLSAVVTTFRHGKPRGIIIVFHGDGAETONGGYKPLMERFAKQYISVSW
DKLVGKSSGNLQSDDRANEVNOVIEMMKVYDPDSTAKIGLMSAQGVVPSKAM
NANNEIDFSILAAPAIWNWQGEYNTCQAKDTGKTKEVQAOQNFLLDSALISKNK
TVESYKQNGKEEMTPGRYNFIRKNNMODATLDSKVTKIYLAELAEKDNVDSAEK
DVTYQIKKENLETQIPNVSHQMINPAIANSELNTNIVGLVAPKPYFLVDNAYLDYCE
QVWSKQ"
/complement(2666..3280)
/gene="lmo1129"
/complement(2666..3280)
/gene="lmo1129"
/note="similar to unknown proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99207.1"
/db_xref="GI:16410545"
/db_xref="SPTREMBL:O8Y720"
/translation="WTLSTKSIDSEKDSIVEVASLVPAGEFKSKFTKTLFKETAHD
ITYACHYIYTEKGENLLIAIDELVCGCLFVTSKSDSKYQKLYQIRKFLFSQCKL
IFLLSLLSHKPVQNERIIDFITVSPNFRQGGIKTLISCHLETDFHDEKTLTVAKKNS
GAYKLYRNIGFOIQENSTLMGLVTGKWKEMWMIQNLVNGK"
/complement(join(3277..3996,4004..4009))
/gene="lmo1130"
/complement(3277..3996)
/gene="lmo1130"
/note="similar to transcription regulators"

```

```
/codon_start=1
/transl_table=11
/protein_id="CAC99208.1"
/db_xref="GI:16410546"
/db_xref="SPTREMBL:O8Y7Y8"
/translation="MULTISUBSTRATE SPECIFICITY OF THE PROTEIN IS DETERMINED BY THE
ADDITIONAL RESIDUES IN THE C-TERMINAL REGION. THE PROTEIN IS A MEMBER OF THE
IHSQEMRINELTTFVAKDDRYFKKVPNQIDHTINYSQAHHNLPDLSEPPYIFSK
QDTGVCLKSIKRRSDYTFPAGYACKSFIAKDEATETQIQLFLEELKYPTTSMILY
ENIYPSLAYPDAMVYTMEDPL"
complement(3288..3293)
/gene="lmo1129"
complement(3288..3293)
/gene="lmo1129"
complement(4004..4009)
/gene="lmo1130"
join(4233..4237,4245..5960)
/gene="lmo1131"
4233..4237
/gene="lmo1131"
4245..5960
/gene="lmo1131"
/notes="similar to ABC transporters, ATP-binding proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99209.1"
/db_xref="GI:16410547"
/db_xref="SPTREMBL:O8Y7Y8"
/translation="MIDKRLPOLVERKSLVLLIFRVLISGLMGLVFAOOLTHVL
EGQNDWMLGTVLVLYGKAILTKLAQEIYQASAEILSLMRRAVMEKAFRLGNE
QGPASTLOVADGELEIYYSRFLPQFYCLIASLIFGLSVFAQFAPOLVLLIC
MPQIPVIAVMAKIRLSGVSDYTNLGRFENLSGLILKAYDQKQDOEIVS
DAERFKATMSLSMQLNSITIMDIISYSCAALIGMSLITMFTNGTISMTGLMFLTY
SAEFFTPMROGLSFLHVMANGISACSLFAYLEKQVYGSVELAKPLKLEIVHLLY
TYGKGAALQDVSVHNGFSYKSGSKSTFVRLNOLPCYQCEIYVNDVPL
ASLSEGAIRKQAVDNDHGYLVANSIRENLLIGNPAANDSLNVLQEQSLADFPVKL
PEQLNLEDGNSLGGQRQLLALRLRQREYVDFEITSGVDSEKILILVQL
ELAKEKIVLIFISHRLYNVLGADQVLDVDFAGKLLVASPERLOQESNYFKNYFLEELAL
LKGG"
5948..5953
/gene="lmo1132"
5948..5953
/gene="lmo1132"
5962..7608
/gene="lmo1132"
5962..7608
/notes="similar to ABC transporters, ATP-binding proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99210.1"
/db_xref="GI:16410548"
/db_xref="SPTREMBL:O8Y7Y8"
/translation="MSEWTLIGWLKFKVLRKMLLAIVLIGLISNLVIMISLIGTY
GIIVLQLOPLNPKYKLFVWACGVVRGLARYLEQLVNDHDIAPRLAIIRERIFATLR
KLPARLSKSGDLVAALITDVEALEVFYHTEISPVFALGTITATVGLFASVDVGL
AIIIVGQIMVGLVPMISYKRNKRIGTAYQAQFVGLNQMVMENTASLQDIFQKIGE
ERLKLTEQERLNQYKRRIGQSELQILGEWVLTATIIILVLSLSQPLETVLI
GTVLSLSFGSLALNALTALLTTFASGRVLYALTEERPVVLFNGQLSDPSEAL
DKVFSHDCKOPLNLSLALPKGWLIGIGSGSGSKSTPLVLLMYWPDQGVKLNN
NELPKTESLHRLEGVMEQRTFLEDITLGNIRLGRKDATLEEVKEAKKAIDTWI
ELTPEGYETIIGQSRNLSGDERGRIGLARFLHDPALLLLDLEPTSNLDYINEQAILLN
TURSEIQDVLVISHRATTLDLAEQLEFENGALKRAAK"
complement(7604..7637,7645..7953,7961..7966))
/gene="lmo1133"
complement(7604..7637)
/gene="lmo1133"
7615..7637
/gene="lmo1132"
7615..7637
/gene="lmo1132"
complement(7645..7953)
/gene="lmo1133"

/notes="similar to B. subtilis Yjcs protein"
/codon_start=1
/transl_table=11
/protein_id="CAC99211.1"
/db_xref="GI:16410549"
/db_xref="SPTREMBL:O8Y7Y6"
/translation="MYLKENTNYLYCTAVIQTTRKYSYEQLEHLEALRAKTYKEPGC
Query Match 29.3%; Score 252.6; DB 1; Length 250050;
Best Local Similarity 59.3%; Pred. No. 8.3e-39;
Matches 429; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 61 ATAAATAATFCAGTAATGGTACTACAGCTAGAGAAGATGTTGTGTTATTCATTAAATGCT 120
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226000 ATAAACGAGCGAAATATCGAACCTAGACTAAAGAGTCCATGTTATATTCATTCAAGCA 226059

QY 121 GGAGGTAAGGCATCCGACCAGTTCGTGTTATTACTACTCATTTCAGTTCACATAAATACCGAG 180
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226060 GCGGGAAAGCGGATTCGTCCAATGCTAGTGTTCGCTACGCTTCAAGCTTTTAAAGTAAAT 226119

QY 181 TATGAGTTAGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACATTAT 240
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226120 CCGCTTTTAGTGTGTAACACCGCAACAGCGTTAGAAATGATTCACATACACGCTTAT 226179

QY 241 CATGATGACCTACCGAGCGATGATAATGATGATTATCGACGAGGAAATTAACAAATCAT 300
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226180 CATGATGATCTACCGAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 226239

QY 301 AAGTATATGCTGAGTGGAGTGGATGCGATATAGCAGGTGATGCTTTTATTAACAAATGCT 360
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226240 AAGTATATGCTGAGTGGAGTGGATGCGATATAGCAGGTGATGCTTTTATTAACAAATGCT 226299

QY 361 GAATTTATTTCAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226300 TCTATTTAGCTGAAACGATATATTTATCTTTTGACACGCGATGCTTTTGATTAACCAA 226359

QY 421 CTGTCATAGCAAGTGGTCAATGTTGGATGGTGGCGGGTCAATGTTAGATATGCAAAAGC 480
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226360 ATTAGTTTATGACGGTGCAGAGGAATGTTGGTGGTCAACTTGCAGACTTGGAGCG 226419

QY 481 GAAGGCCAACCAATTCATCTTGAACCTTTGGAATGATACACAAACAAACAGGAGCA 540
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226420 GAAACAAACAGTACGCTAGAGAGATTATCATCCATTCATGCAGAAACAGGGGTGA 226479

QY 541 TTATTAACCTTTTGGGTTATGAGTCAGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226480 TTATTAATTTATGCTTAACCTCTGCAGCAAAATTTGCGGAACCTGATCCAGAACAC 226539

QY 601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAGATGATTTA 660
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226540 AAACGCTTACGAATTTTTCAGAGAATATTTGGGATTTGGATTTCAAAATTAGCGAC 226599

QY 661 TTAGACTGCTATGCTGATGAAGCAAAAGTTAGGTAAACAAAGTGGCAGCGATCTTGAAT 720
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226600 TTAGATGTAATTTGGTATGAACGAAATGGGTAAACAGACAGGCGCGGCGCTTTCTG 226659

QY 721 AATAAAAGTACCTGCTGATGTTATTTAGGGAAGATGGGCGAGCAAGATAAATTTGACTTAT 780
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226660 AATAAAAGTACCTGCTGATGTTATTTAGGGAAGATGGGCGAGCAAGATAAATTTGACTTAT 226719

QY 781 CAT 783
|||||
Db 226720 CAT 226722

RESULT 11
AX416832 LOCUS
AX416832 DEFINITION Sequence 5635 bp DNA linear PAT 14-JUN-2002
AX416832 ACCESSION
AX416832.1 VERSION
AX416832.1 GI:21449341
AX416832.1 KEYWORDS
AX416832.1 Listeria monocytogenes ATCC 19115.
```


LOCUS DEFINITION	AB003187	1080 bp	DNA	linear	BCT 25-MAR-1999
Micrococcus luteus DNA for farnesyl diphosphate synthase, complete cds.					
ACCESSION	AB003187	GI:2982676			
VERSION	AB003187.1				
KEYWORDS	farnesyl diphosphate synthase.				
SOURCE	Micrococcus luteus (strain:B-P 26)	DNA.			
ORGANISM	Micrococcus luteus				
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococccineae; Micrococcaceae; Micrococcus.				
AUTHORS	Shimizu, N., Koyama, T. and Ogura, K.				
TITLE	Molecular cloning, expression, and characterization of the genes encoding the two essential protein components of Micrococcus luteus B-P 26 hexaprenyl diphosphate synthase				
JOURNAL	J. Bacteriol. 180 (6), 1578-1581 (1998)				
MEDLINE	98175686				
REFERENCE	2 (bases 1 to 1080)				
AUTHORS	Koyama, T.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-APR-1997) Tanetoshi koyama, Tohoku University, Institute for Chemical Reaction Science, Katahira 2-1-1, Aoba-ku, Sendai, Miyagi 980-77, Japan (E-mail: koyama@icrs.tohoku.ac.jp, Tel: +81-22-217-5621, Fax: +81-22-217-5620)				
FEATURES	Location/Qualifiers				
source	1..1080				
	/organism="Micrococcus luteus"				
	/strain="B-P 26"				
	/db_xref="taxon:1270"				
gene	105..980				
	/gene="fps"				
CDS	105..980				
	/gene="fps"				
	/codon_start=1				
	/transl_table=11				
	/product="farnesyl diphosphate synthase"				
	/protein_id="BAA25265.1"				
	/db_xref="GI:2982677"				
	/translation="MLQPKLTMRDFNLNLSLNKYNKHPAOSRLHEAINYSLSAGK RIRPLVLTLDLSLGGNAHDCGLPFCIALEMLHTVSLIHDLPAMDNDYDPRGLTWHK RFDEATLILAGDALLTDFQCLATQINAEIKLSLNLSTAGSGNKNVYQMLDMQGG EHKUTLNELEIRHIHKTGELIRAAIVSAGLIMFNDAEQNLINIGKNVYQMLDMQGG DILDVEGFSFGNITGVSDLNDRKSTVSLGLEAKQLLNKLTETDYDLKTLQPIN DNKLTLYIVYVERNK"				
BASE COUNT	377 a 161 c 198 g 344 t				
ORIGIN					
Query Match	27.1%	Score 233.4	DB 1	Length 1080	
Best Local Similarity	57.4%	Pred. No. 9.6e-35			
Matches	461	Conservative 0	Mismatches 336	Indels 6	Gaps 2
QY	6	GAATCTCCGATGAATAAATTAATAGATGAAGTCAATTAATGAATTAATTCGGTTCGGATAAA	65		
Db	116	GAACGTGACATGAATTCGTCGATCTTTCTTAATCTGATTAATGAATCAATGTTTAAATNAATA	175		
QY	66	TAAATCAGTAATGGATCTACGCTAGAGAAAGATGTTGTTATTTCATTAAATGCTGGAGG	125		
Db	176	TCATCCTG---CGCAATCAAGACTACATGAAGCAATCAACTATTCATTAACTGCGAGG	232		
QY	126	TAAACGATCCGACAGTTCGTTTATTACTCATTAGATTCACTTAATACCGAGTATGA	185		
Db	233	CAAGCGCATTAGACCACTACTTGTCTGACTACTTTAGATAGTTAGTGGCAATGCACA	292		
QY	186	CTTAGGTATGAGACGCCAATTCGACTAGAAATGATTCATACATATTCACCTATTTCATGA	245		
Db	293	TGACGCTTTACCATTTGGCATTCGCGCTTGAATGATTCATACGATTTCTTTAAATTCACGA	352		
QY	246	TGACCTACCAGCGATGATGATGATTATCGACGAGGAGAAATTAACAAATCATATAAGT	305		
Db	353	TGACTTGGCGCAATGGATATGATGACTATCGTCGCGGTAACTACGAAATCATAGCG	412		
QY	306	ATATGGTGAGTGGACTGGCATATTAGCAGGTGATCTTTTATTAACTAAAGCATTTGAAT	365		

RESULT 12
AB003187

```
Db 413 TTTTGTGAAGCAACAGCTATACCTCGCTGAGATGCAATGCTCCTCAGTATGCTTTTCAATG 472
QY 366 TATTTCAAGTGATGATAGATTAAGTGAAGTAAATAAATAAAGTTCTCAACGCGTGTG 425
Db 473 CATTTTAA---ATACGAGTTAAACGACAGAAATTAATATCATTCATTATTAATTAAG 529
QY 426 AATAGCAAGTGGTCATGTTGGAATGTTGGCGGTCAAAATGTTAGATGATGCAAGCGAAGG 485
Db 530 TACTGCTTCTGGATCTAATGGCATGTTTACGGCCAAATGCTCGATATGCAAGGTGAACA 589
QY 486 CCAACCAATTTGATCTTTGAAATTTGGAATGATACACAAAACAAAACAGGAGGACATTAAT 545
Db 590 TAAACATTTGACATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 649
QY 546 AACTTTTGGCGTTATGAGTGCAGCAGATATCGCTAATGATGATGATGATGATGATGATGAT 605
Db 650 TCGTGCAGCAATTTGTAAGTGCAGGTATCATATGAATTTTAATGATGCAAAATTCAGCA 709
QY 606 TTTGAAAGTTATAGTTTATCATTATGATGATGATGATGATGATGATGATGATGATGATGAT 665
Db 710 ACTCAATATATTTGTTGAAGACGTAGTTTAAATGTTTCCAGATAAAGATGATATTTCTGA 769
QY 666 CTGCTATGTTGATGAAGCAAGTTAGTAAATAAGTGGCAGCGATCTTTCAGAAATATAA 725
Db 770 TGTGAAGGTGCTTCGAAATATCGGAAACAGTTGGCAGTGAATTAATTAATGATAA 829
QY 726 AAGTACGTACGTGAGTTTATAGGAAAGATGCGCAGAGATGATGATGATGATGATGATGATGAT 785
Db 830 AAGTACATATGATCACTCTTAGGACTTGAAGCTTCAAGCTTCAAGCAATTAATTAATGATGAT 889
QY 786 AGACGAGCAGTGGATGACTAA 808
Db 890 AACTGAACATATGATGCGCTTA 912

RESULT 13
AX417035/c
LOCUS 195269 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 4026 from Patent WO0228891.
ACCESSION AX417035
VERSION AX417035.1 GI:21449645
KEYWORDS
SOURCE Listeria innocua.
ORGANISM Listeria innocua
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE 1
AUTHORS Glaser, P. and Kunst, F.
TITLE Listeria innocua, genome and applications
JOURNAL Patent: WO 0228891-A 4026 11-APR-2002;
Pasteur Institut (FR)
FEATURES
    Location/Qualifiers
        source
            1..195269
                /organism="Listeria innocua"
                /db_xref="taxon:1642"
                /note="second part of seq 8-first part = seq 8 (000001 - 349980)-second part = seq 4026 (300001 - 495269)"
BASE COUNT 57412 a 39214 c 31781 g 66862 t
ORIGIN

Query Match 26.9%; Score 231.2; DB 6; Length 195269;
Best Local Similarity 56.4%; Pred. No. 1e-34;
Matches 431; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 20 ATAAATTAATAGATGAAGTCAATATGAATTTATCGGTTCCGATATAATAATCAGTAATGG 79
Db 66205 ATCAATATAAAGTACTTGATGAGTCGCTTTTAAAGAAATAATAATGCCAATATCG 66146
QY 80 ATATCAGCTAGAGAAAGTATGTTGATATCAATTAATGCTGGAGTAAACGCATCCGAC 139
Db 66145 AACCTAAGTTGAAGAAATCAATGTTATATTCAGTCAAGCTGGCGGAAACGAATTCGTC 66086
QY 140 CAGTTCTGTATTAATCACTTATAGTTTCAATTAATACCGAGTATGATGATGATGATGATGAT 199
```

```
Db 66085 CAATGCTGTTTTCGCAACGCTTCAAGCCCTTAATAATATGAGCCAATATGCGCGGTTTAAAAA 66026
QY 200 GCCCAATTCGACATAGAAATGATTATCATACATATTTCACTTATTATGATGACTACACGCGA 259
Db 66025 CAGCTACGCGACCTTGAATGATTCATACGCTTAAATTCAGATGATATTTACGAGCAA 65966
QY 260 TGGATATGATGATATTCGACGAGGAAATTAACAATATCAATAAGTATATGTTGAGTGGGA 319
Db 65965 TGGATGATGATGATATTCGACGCGGAAATGAGCAACCAATAAATAATATACGGTGTGTCAA 65906
QY 320 CTCGCGATATTCAGCGTGTGCTTTTAACTAAAGCATTTGAACATTTTCAAGTGTATG 379
Db 65905 CAGCAATTTAGCAGAGATGCTTTTGTAACTCTGCTTTTCTATTTAGCTGNAAGATG 65846
QY 380 ATAGATTAATCTGATGAAGTAAATAAATAAAGTTCTACAACGCGTGTCAATAGCAAGTGGTC 439
Db 65845 AATATCTATCTTTCGAAACGCGTATTTGCTTTAATTAATCAATTTAGTTATAGCAGTGGAG 65786
QY 440 ATGTTTGAATGTTGCGCGGTCAAAATGTTAGATATGCAAAAGCGAGGCCAACCAATTTGNTC 499
Db 65785 CAGAAGGCATGTTAGTGGCAACAACCGCATATGGAAGCTGAAATAAACAAGTCACAT 65726
QY 500 TTCAAACTTTGGAAATGATACACAAAACAAAACAGGAGCATTTTAACTTTTTCGCGTTA 559
Db 65725 TAGAAGATAGATCAATCCACGCTCGCAAACTGCTGAACATTTAATTTTCGCAGTAA 65666
QY 560 TGAGTGCACGATATCGCTTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATA 619
Db 65665 CCTCAGCCGCAAAATTCGTTGAGCAACTCCAGAACAAACAACCAATTTAGCAATTTTGTG 65606
QY 620 GTTATCATTTAGTATGATGTTTCCAGATTAAGATGATTTATTTAGACTGCTATGTTGATG 679
Db 65605 CAGAAAATATCGGCATTTGTTTCAATTTAGCGACCATATTTTATGATGTAATTTGCGCATG 65546
QY 680 AAGCAAGTTAGTAAATAAGTGGCAGCGATCTTCAAAATAATAAAGTACGTACGTGA 739
Db 65545 AAACAAAATTTGGTAAAGACAGCGGTCTGCGCTTTTCTGTAATAAAGTACCTATCCCG 65486
QY 740 GTTTATTAGGAAAGATGCGCAGAGATGAAATTTGACTTTATCAT 783
Db 65485 GATTACTCACATCGGAAGTGTAAACGGCATTAATGAGCAT 65442
```

RESULT 14

```
AL596168
LOCUS 333050 bp DNA linear BCT 06-JUN-2002
DEFINITION Listeria innocua Clp11262 complete genome, segment 6/12.
ACCESSION AL596168 AL592022
VERSION AL596168.1 GI:16413677
KEYWORDS
SOURCE Listeria innocua.
ORGANISM Listeria innocua
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE 1
AUTHORS Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetoui, F., Couve, E., de Daruvar, A., Deboux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dusserre, O., Entian, K. D., Fsihi, H., Portillo, F. G., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Krest, J., Kuhn, M., Kunst, F., Kurapkaj, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablo, B., Perez-Diaz, J. C., Purcell, R., Rammel, B., Rose, M., Schlueter, T., Simoes, N., Tiersch, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. and Cossart, P.
Comparative genomics of Listeria species
Science 294 (5543), 849-852 (2001)
TITLE Listeria innocua
JOURNAL Science 294 (5543), 849-852 (2001)
MEDLINE 21537279
PUBMED 11679669
REFERENCE 2 (bases 1 to 333050)
AUTHORS Glaser, P., Frangeul, L. and Rusniok, C.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Glaser P., Institut Pasteur, Genomique des
```

Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cedex 15, FRANCE
E-mail: pglaser@pasteur.fr
Phone: +33 (0)1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
Location/Qualifiers
1. 333050
/organism="Listeria innocua"
/strain="Clp11262"
/db_xref="taxon:1642"
915..4679
/gene="lin1204"
915..920
/gene="lin1204"
927..4440
/gene="lin1204"
/note="similar to internalin proteins (LPXTG motif)"
/codon_start=1
/transl_table=11
/protein_id="CAC96435.1"
/db_xref="GI:16413678"
/db_xref="SPTREMBL:Q92CG8"
/translation="MKRKISSIVVIGIMFQSLTWYPTITEARENEQKEEINKPKSKIT
KGLTNSLKYTKTILETDYDSVFPDSALAKVVAKEATGSENTQLVTAQDLNKLKSL
NGYKNGISVLTGIDLVNVTISLNNQVTDISPIDQPNLVSLSVKKNQISSLLNA
QNLKPLATTIDENPDNLTIDIQPOLVDVKTSGYGLKLTIVIAKNPELVNLG
QYTIKRVFSQVASTKVELYNLPKRVNLRNSINELKVTDDLAIEDLPLENELTD
TVFDNIQNLPTKLSKQNLQEEVLDKTDVENPNLMTDIOQNLAKILINVOQDP
QLVDVKTSDYKELSAITVIKKNPELVNLQYTIKRVFSQVASTKVELYNLPKVR
KYNLRNSINELKVTDDLAIEDLPLENELTDVFDNIQNLPTKLSKQNLQEEVLD
DKTDVENPNLMTDIOQNLAKILINVOQDPOLVDVKTSDYKELSAITVIKKNPEL
VNLGPIKMONVFSQVASTKVELYNLPKRVNLRNSINQIELNLSVVKVDNLT
NKITNDSIEKFRGMFILATNLNKNQIININMLDGFQPTILNIDNLSISVLPSNLT
MKPKLSRSALNQTGLKAIIVDDSDLIINNEISNFKLETPSPNFGTYANEKIT
WSSERIKNLTVEFSQSELINVTIDGTFSKVTQFFKRTTPVINADSEIHPQOTE
KTEAEFLDQAQYTDLLAIDSFETMNLKVGKYTVTLNENLDGIKATPEYIVY
IDAVGANITVYDESEKLAENSLTGNVEEYSSSAKEITLGYTLTEMPNAGOE
SLEQTVTVYISKVNPAPKLTVOYTDDEGMELAPETLSGVDENVYTTAKFTGYE
LIETPSNAGFSENAQIVTVYVRAIKAPILAKETVYVQDELGAKISETVLAGEI
GETYTTATIDGYTLIKSPTNANGIFNENPQTVTVYVQQLNSPAAKITVVKHLDEN
NELAPSEVLGNDQRAYTKPKREIKDYSLVKVPTNASGKFTTEATQTVIYHKKNISPT
SSYITVKYVDETGRELAISKVLNGINDSYATTAKEIKGYTLVQKPNATGKFTQNO
TIKYVYRANKAEVNPNSDVLGNSNGDIOPTNDIKIPNDVQPTNEAPQTLPKTGSO
PANLJFGGILLVLLSTQWLHNKRRKAKSHNC"
4659..4679
/gene="lin1204"
Join(4883..6241,6249..6276)
/gene="lin1205"
4883..6241
/gene="lin1205"
/codon_start=1
/transl_table=11
/protein_id="CAC96436.1"
/db_xref="GI:16413679"
/db_xref="SPTREMBL:Q92CG7"
/translation="MKMLPESKVESIRKEGFLNRAVEAYREFYPTVSNVSNFKALHD
LGITENHDFIQLTTPDLNLVFNQSDTFCIGTNGTEGVPVIELPLGAIVGADDIN
FKFITNMLGTDEEGKAKLYLPNNYDIDPEGYIVRKPSSFRFLICLRAMVHQASD
YEKAFELTKTVKFLPEKAKNPTSTFHDGSHRKAISTPYVEGFDYWEVTKWALDN
DETDPYQTMQYGLLAKIIGLAPNKEPNPDSDTKALLIEAAEKADMMFVNSVTDPPAA
IIVPKGNMWEAYVNDNFYEKTYLNLVPRWERFYQATLETMMFMHKVFGSGVYMLG
VRDKESGILDGKSYTLKVPVPTPIFSWSTVYEMDRSEIVTEQFPALNSIKDTP
VADAGNTTLFGPNPPEDESFLWQTVDPDANWFTIYRIYGPTEFAFONSWQLDYFEE
VK"
complement(join(6249..6276,6283..6876,6885..6890))
/gene="lin1206"
6249..6276
/gene="lin1205"
complement(6249..6276)
/gene="lin1206"
complement(6283..6876)
/gene="lin1206"
/note="similar to B. subtilis ydeI protein"
/codon_start=1

/transl_table=11
/protein_id="CAC96437.1"
/db_xref="GI:16413680"
/db_xref="SPTREMBL:Q92CG6"
/translation="MAKTELNPKVDLSFNKPKSTWQAEKALREMAITFLEEFKWKG
PCYALNGSVFLIHGFKNYKCALLEGALLRDPDNLVQOTENVOAARQIRFTNLQEI
LDQKELKYIONAEVEKAGLEVALKPAETPIPELLVFKFEELPALATAFALTTPG
RQKAYLLYFAAPKQSKTRVSRIEKYEATILDLGLGUND"
complement(6885..6890)
/gene="lin1206"
complement(6994..7414)
/gene="lin1207"
complement(6994..7401)
/gene="lin1207"
/codon_start=1
/transl_table=11
/protein_id="CAC96438.1"
/db_xref="GI:16413681"
/db_xref="SPTREMBL:Q92CG5"
/translation="MTLNVLNFRQSRDAIAFYEIFGTCTDMLTYGEIESAEPI
EDSLKDLVNASLVLDGVKVMFSDVPKSMPLTFGDNITLVITDITDEIKLTKQFHLAE
GGNIIMPGLKTFWSEKYQVTDKFGIGWQLNLS"
complement(7409..7414)
/gene="lin1207"
7553..8165
/gene="lin1208"
7553..7558
/gene="lin1208"
7566..8165
/gene="lin1208"
/note="weakly similar to phosphoglycerate mutase 1"
/codon_start=1
/transl_table=11
/protein_id="CAC96439.1"
/db_xref="GI:16413682"
/db_xref="SPTREMBL:Q92CG4"
/translation="MKKTYLMRHGQTLFNQKKTQGFCDAPLTDLGITQAKIAGSYF
KENNITFDVYSSTSERACDTLELITDKSYQRLKLNWNGFAGFEGESDLNPLPYG
DFPAGYGERVDYFRDLVYTMERIMSNQENHDTVLAVSHGAACAQAPARYWEKTSKIG
VTGLNCCILKPEYENGRTLVNFINHDFENGTHMESVR"
Join(8347..8352,8361..8723)
/gene="lin1209"
8347..8352
/gene="lin1209"
8361..8723
/gene="lin1209"
/codon_start=1
/transl_table=11
/protein_id="CAC96440.1"
/db_xref="GI:16413683"
/db_xref="SPTREMBL:Q92CG3"
/translation="MSKYDAEIVNAQOQSITGOQEKIRRLKELITTKTERGEOITSSIP
ROIDLPPDFWQCKSDSVMEVLQRRLOFNQNNQNAVVDLVQELRAAMTKMNOISDYQ
ADISYTNLKOMEEVNV"
8701..8706
/gene="lin1210"
8701..8706
/gene="lin1210"
8716..8991
/gene="lin1210"
8716..8991
/gene="lin1210"
/codon_start=1
/transl_table=11
/protein_id="CAC96441.1"
/db_xref="GI:16413684"
/db_xref="SPTREMBL:Q92CG2"
/translation="MSDIQVNESQIIKSNQFAGKQVLYKKSSSTSHNNKLATYHKI
QTILTKMDTQMNLQQLGERTTQDIEKNAEAFVNLKLLTDIEVTG"
8992..10532
/gene="lin1211"
8992..8997
/gene="lin1211"

```
CDS      8999..10459
/gene="lin1211"
/codon_start=1
/transl_table=11
/protein_id="CAC96442.1"
/db_xref="GI:16413685"
/db_xref="SPRMBL:092CG1"
/transl_table="MKIDVRELNEIYSENIVMLKKEYENLINIAMNATTNFTVETQBF
KGTADASRAYLQEAIVKPVQOKALEVNLKMTKILSEVIADEAEPFGANGMIDIOSIEM
EYKRSNOVSDDMBRVDLNNLIOBANETYPQSSNLGFFELHDAITETINRMK
LEDFENKMWELSKLOKELDMLTQVNDKILITSVQACGLPFGKPKSRDEDT
VQBYQYVAVANAKABEILVDGQRIEFTDFTYENKNEVNGTLVNEVNGEMI
LFVQRKDIESAVSPGVFPDFPFLANGSETDGTIKLYHKRSKDIPEGLSPVNEEL
KAWADFFADTQNSYDAGMAYLKAHNKPEKLSFTDLDKVTATNWSGKNRFDAG
EVEAHADMLTDSILTALAAMPDKIDRPGNTDEKQGYVHAASIAQMDQKEDGMNKK
SPNSIWGSGGYSAVLQNLHMQENTPKAKRSRPRQ"
terminator
/gene="lin1211"
join(10896..11450,11456..11489)
/gene="lin1212"
10896..11450
/gene="lin1212"
/codon_start=1
/transl_table=11
/protein_id="CAC96443.1"
/db_xref="GI:16413686"

Query Match      26.9%; Score 231.2; DB 1; Length 333050;
Best Local Similarity 56.4%; Pred. No. 9.1e-35;
Matches 431; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 20 ATAAATTAATAGATGAAGTCAATTAATGAATATATCGGTGCGATAAATAATCAATAGTATGG 79
DB 174746 ATCATTAATAAAGTACTTGATGAGTCGCTTTTAAAGAAATAAATATCGCAATATCG 174805
QY 80 ATACTCAGCTAGAGAAAGTATGTTGATTCATTAATGCTGAGGTAACGCGATCCGAC 139
DB 174806 AACCTAAGTTGAAGAATCAATGTTATTCAGTGAAGCTGCGGAAACGAAATTCGTC 174865
QY 140 CAGTTCCTGTTATTCACCTTTAGATTCATTAATACCAAGTATGAGTTAGTGAAGA 199
DB 174866 CAATGCTGTTTTCGAACGCGTATTCGTTTAAATCAAAATAGTATAGCAGTGGAG 175165
QY 320 CTGCGATATTAGCAGGTGATGCTTTTAACTAAAGCATTTGAACCTTATTTCAAGTGATG 379
DB 175046 CAGCAATTTAGCAGGAGATGCTTTGTTAACTACCTCGCTTTTCTATTTAGCTGNAAGTG 175105
QY 380 ATAGATTAAGTGAAGTAAATAAAGTCTTACACGCGTGCTCAATAGCAAGTGGTC 439
DB 175106 AAAATCTATCTTCGAACGCGTATTCGTTTAAATCAAAATAGTATAGCAGTGGAG 175165
QY 440 ATCTTGGAATGTCGCGGTGCAATGTTAGATATGCAAGCGAAGCGCAACCAATTTGATC 499
DB 175166 CAGAAGCGATGTTAGTGGCAACAGCGCATATGGAAGCTGAAATAACCAAGTCAAT 175225
QY 500 TTGAAATTTGGAAATGATACACAAACAAACAGGAGCATTTATTAACCTTTTTCGCGTTA 559
DB 175226 TAGAAGCAATAGCATCAATCCAGCTCGCAAAACTGCTGNACTATTAATTTTCGAGTAA 175285
QY 560 TGAGTCCAGCATATCGCTATGTCGATGATACAACTAAAGCAATTTAGAAAGTTATA 619
DB 175286 CCTCAGCCCGCAAAATCGCTGAGCAACTCCAGAACAAACAAACGATTTACGAAATTTTG 175345
QY 620 GTTATCATTTAGTATGTTCCAGATTAAGATGATTTATAGATCTGCTATGTTGATG 679

Db 175346 CAGAAAATATCGCATTTGGTTTTCAAAATAGCGACGATATTTAGATGTAATTTAGCGCATG 175405
QY 680 AAGCAAGTTAGTTAAAAAAGTGGCGAGCGATCTTGAATAATAATAAAGTACGTACCTGTA 739
Db 175406 AAACAAAATGCGTAAAGACACAGGGTGCACGCTTTTCTGCAATAAAGTACCTATCCCG 175465
QY 740 GTTTATTAGGAAGATGGCGACGAAGATAAATTTGACTTTATCAT 783
Db 175466 GATTACTCACATCGAAGGTGCTAAACGGCGCATTAATATGACAT 175509

RESULT 15
LOCUS AX417044 349980 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 4035 from Patent WO0228891.
ACCESSION AX417044
VERSION AX417044.1 GI:21449654
KEYWORDS
SOURCE Listeria innocua.
ORGANISM Listeria innocua
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE
1 Glaser, P. and Kunst, F.
AUTHORS Listeria innocua, genome and applications
TITLE Patent: WO 0228891-A 4035 11-APR-2002;
JOURNAL Pasteur Institut (Fr)
FEATURES
source
1..349980
/organism="Listeria innocua"
/db_xref="taxon:1642"
/notes="seq 2058, original length: 3,011,208 replaced
by-seq 2058: 0,000,001 to 0,349,980-seq 4032: 0,300,001 to
0,649,980-seq 4033: 0,600,001 to 0,949,980-seq 4034:
0,900,001 to 1,249,980-seq 4035: 1,200,001 to
1,549,980-seq 4036: 1,500,001 to 1,849,980-seq 4037:
1,800,001 to 2,149,980-seq 4038: 2,100,001 to
2,449,980-seq 4039: 2,400,001 to 2,749,980-seq 4040:
2,700,001 to 3,049,980-seq 4041: 3,000,001 to 3,011,208"
BASE COUNT 116021 a 60232 c 66706 g 107021 t
ORIGIN

Query Match      26.9%; Score 231.2; DB 6; Length 349980;
Best Local Similarity 56.4%; Pred. No. 9e-35;
Matches 431; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 20 ATAAATTAATAGATGAAGTCAATTAATGAATATATCGGTGCGATAAATAATCAATAGTATGG 79
DB 192746 ATCATTAATAAAGTACTTGATGAGTCGCTTTTAAAGAAATAAATATCGCAATATCG 192805
QY 80 ATACTCAGCTAGAGAAAGTATGTTGATTCATTAATGCTGGAGGTAAACGCGATCCGAC 139
DB 192806 AACCTAAGTTGAAGAATCAATGTTATTCAGTGAAGCTGCGGAAACGAAATTCGTC 192865
QY 140 CAGTTCCTGTTATTCACCTTTAGATTCATTAATACCGAGTATGAGTTAGTATGAAGA 199
DB 192866 CAATGCTGTTTTCGAACGCGTTCGAACGCTTAAATATTCAGCAATTCGCGGTTAAAAA 192925
QY 200 GCGCAATTCGCTAGAAATGATTACATATTCATCTATTTCATGATGACCTACCGACGCA 259
DB 192926 CAGCTACGCGACTTGAATATGATTCATACGCTAATTCACGATGATTTACCGACAA 192985
QY 260 TGGATTAATGATGATTATCGACGAGGAAATAATCAAAATCATAAAGTATATGTTGAGTGA 319
DB 192986 TGGATGATGATGACTATCGACGCGGAAATGGACGAACCATAAATATACGGTGTGCA 193045
QY 320 CTGCGATATTAGCAGGTGATGCTTTTATTAACCTAAACGATTTGAACCTTTTCAAGTGATG 379
DB 193046 CAGCAATTTAGCAGGAGATGCTTTTAACTACCTCGCTTTTCTATTTAGCTGTAAGATG 193105
QY 380 ATAGATTAAGTGAAGTAAATAAAGTCTTACACGCGTGCTCAATAGCAAGTGGTC 439
DB 193106 AAAATCTATCTTCGAACGCGTATTCGTTTAAATCAAAATAGTATAGCAGTGGAG 193165
```

QY 440 ATGTTGGAATGGTCGGCTCAATGTTAGATATCCAAAGCCAAAGCCCAACCAATTGATC 499
 Db 193166 CAGAAGGCATGGTAGGTGGCAACAAGCCGATATGGAAGCTGAAAAATAAACAAAGTCACAT 193225
 QY 500 TTGAAACTTTGGAAATGATACAAACAAACAAAGGAGCATTATTAACTTTTGGCGTTA 559
 Db 193226 TAGAAGAAGTACATCAATCCAGCTCCCAAACTGGTGAACCTATTAAATTTTCGCAGTAA 193285
 QY 560 TGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACAATTTAGAAAGTTATA 619
 Db 193286 CUTCAGCGGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAACGATTACGAATTTTG 193345
 QY 620 GTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTATTAGACTGCTATGCTGATG 679
 Db 193346 CAGAAAATATCGGCATTGGTTTCAAAATAGCGAGATATTTTAGATGTAAATTTGGCGATG 193405
 QY 680 AAGCAAGTTAGGTAAAAAAGTGGCAGCGATCTTGAATAATAAAGTACGTACGTGA 739
 Db 193406 AAACAAAAATGGTAAAAAGACAGGGTTCGACGCTTTTCTGAATAAAGTACCTATCCCG 193465
 QY 740 GTTTATTAGGNAAGATGGCGCAGAGATAAATTGACTTAATCAT 783
 Db 193466 GATTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCAT 193509

Search completed: May 30, 2003, 16:56:52
 Job time : 2487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 13:05:08 ; Search time 280 Seconds

(without alignments)
6924.895 Million cell updates/sec

Title: US-09-925-637-63

Perfect score: 861

Sequence: 1 atgacgaatcaccgatgaa.....tattagaatcgttgattta 861

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	100.0	861	22	AA500821 S. aureus HGS072 e
2	853	99.1	882	21	AA92031 Staphylococcus aur
3	846.6	98.3	882	23	AA54848 Staphylococcus aur
4	831.6	96.6	864	23	AA51602 Staphylococcus aur
c 5	807	93.7	1893	18	AAV74466 Staphylococcus aur
6	449.8	52.2	909	24	ABN90871 Staphylococcus epi
7	449.8	52.2	3444	22	AAH54780 S. epidermidis gen
c 8	449.8	52.2	4045	22	AAH54565 S. epidermidis gen
c 9	413	48.0	413	23	AA550245 Staphylococcus aur

c 10	406.8	47.2	783	22	AAH52415 S. epidermidis ope
11	337	39.1	337	23	AA50698 Staphylococcus aur
12	252.6	29.3	2944528	24	ABA03041 Listeria monocytog
13	239.8	27.9	5635	24	ABQ71010 Listeria monocytog
c 14	231.2	26.9	495269	24	ABQ67195 Listeria innocua c
15	231.2	26.9	3011208	24	ABQ69245 Listeria innocua d
c 16	204	23.7	3454	22	AAH54634 S. epidermidis gen
17	196.8	22.9	882	23	AA52896 Enterococcus faeca
18	195	22.6	7528	20	AAI12992 Enterococcus faeca
19	175.2	20.3	870	24	ABN66185 Streptococcus poly
20	174.8	20.3	2365589	24	ABA90521 Genomic sequence o
21	164	19.0	811	24	ABK73724 Bacillus lichenifo
22	157.2	18.3	894	14	AAQ39243 FPS DNA. Bacillus
23	157.2	18.3	894	17	AAI40229 Native farnesylidip
24	155.8	18.1	1337	20	AAI77873 L. monocytogenes c
25	155.6	18.1	894	17	AAI40226 Mutant farnesylidip
26	155.6	18.1	894	17	AAI40227 Mutant farnesylidip
27	155.6	18.1	894	17	AAI40228 Mutant farnesylidip
28	155.6	18.1	894	19	AAV38455 DNA encoding farne
29	155.6	18.1	894	19	AAV18516 Bacillus stearothe
30	155.6	18.1	894	20	AAI86781 Farnesyl diphospha
31	155.6	18.1	894	21	AAI63809 Bacillus stearothe
32	155.6	18.1	894	24	ABK96797 B. stearothermophi
33	154	17.9	894	17	AAI40225 Mutant farnesylidip
34	154	17.9	894	20	AAI86780 Geranyl diphosphat
35	154	17.9	894	24	ABK96808 B. stearothermophi
36	153.6	17.8	873	24	ABN66184 Streptococcus poly
37	153.6	17.8	2155561	24	ABN71527 Probe B500 for Mic
38	150.6	17.5	486	19	AAV15824 Methanococcus jann
39	137	15.9	1664976	19	AAV21209 Taxus cuspidata ge
40	136.4	15.8	885	21	AAI13985 Taxus GGPP synthas
41	134.8	15.7	1179	21	AAI13994 Taxus GGPP synthas
42	134.8	15.7	1179	21	AAI13995 Taxus GGPP synthas
43	134.8	15.7	1179	21	AAI13996 Taxus GGPP synthas
44	134.8	15.7	1179	21	AAI13997 Taxus GGPP synthas
45	134.8	15.7	1179	21	AAI13998 Taxus GGPP synthas

ALIGNMENTS

RESULT 1

AA500821

ID AA500821 standard; DNA; 861 BP.

XX

AC AA500821;

XX

DT 04-JUL-2001 (first entry)

XX

DE S. aureus HGS072 encoding Farnesyl diphosphatesynthase, IspA.

XX

KW Farnesyl diphosphatesynthase; IspA; immunogen; vaccine; antibody;

KW wound infection; cellulitis; burn infection; eyelid infection;

KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;

KW skin infection; scalded skin syndrome; toxic epidermal necrosis;

KW Ritter's disease; Lyle's disease; toxic shock syndrome; endocarditis;

KW HGS072; ds.

XX

OS Staphylococcus aureus.

XX

FH Key Location/Qualifiers

FT CDS 1..861

FT

FT /tag= a

FT /product= "IspA"

FT /partial

FT /note= "No stop codon"

WT WO200116292-A2.

XX

PD 08-MAR-2001.

XX

PF 31-AUG-2000; 2000WO-US23773.

XX

CC tracheitis, acute epiglottitis, thyroiditis, empyema and lung abscesses),
 CC cardiac infections such as infective endocarditis, gastrointestinal
 CC infections including secretory diarrhoea, splenic abscesses and
 CC retroperitoneal abscesses, CNS infections such as cerebral abscesses, eye
 CC infections (including blepharitis, conjunctivitis, keratitis,
 CC endophthalmitis, preseptal and orbital cellulitis and dacryocystitis),
 CC kidney and urinary tract infections such as epididymitis, intrarenal and
 CC perinephric abscesses and toxic shock syndrome, skin diseases (including
 CC impetigo, folliculitis, cutaneous abscesses, wound infection and
 CC bacterial myositis), bone and joint infections such as septic
 CC arthritis and osteomyelitis, septic thrombophlebitis, food poisoning and
 CC scalded skin syndrome. In addition, they can be used to treat
 CC diseases caused by helicobacter pylori, including stomach cancer, stomach
 CC ulcers and gastritis.

XX Sequence 882 BP; 328 A; 118 C; 186 G; 250 T; 0 other;

Query Match 99.1%; Score 853; DB 21; Length 882;

Best Local Similarity 99.4%; Pred. No. 2, 9e-192;

Matches 856; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1	ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATTAATGAATTCGGTTCGG	60
1	ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATTAATGAATTCGGTTCGG	60
61	ATAAATAATCAGTAATGGATCTACCTAGAGAAAGTATGTTGTTATTAATGCT	120
61	ATAAATAATCAGTAATGGATCTACCTAGAGAAAGTATGTTGTTATTAATGCT	120
121	GGAGTAAACGATCCGACCTCTCTGTTATTAATCTACCTAGATTCACCTAAATACCCAG	180
121	GGAGTAAACGATCCGACCTCTCTGTTATTAATCTACCTAGATTCACCTAAATACCCAG	180
181	TATGATTTAGTATGAAGACGCAATTCACCTAGAAATGATTCATATTCATTTAT	240
181	TATGATTTAGTATGAAGACGCAATTCACCTAGAAATGATTCATATTCATTTAT	240
241	CATGATGACCTACCGATGATTAATGATGATGATGATGATGATGATGATGATGATGAT	300
241	CATGATGACCTACCGATGATTAATGATGATGATGATGATGATGATGATGATGATGAT	300
301	AAAGTATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	360
301	AAAGTATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	360
361	GAACCTATTTCAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	420
361	GAACCTATTTCAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	420
421	CTGTCATAGCAAGTGGTCATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTT	480
421	CTGTCATAGCAAGTGGTCATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTT	480
481	GAAGCCCAACCAATTTGAACTTTGAAATGATGATGATGATGATGATGATGATGATGAT	540
481	GAAGCCCAACCAATTTGAACTTTGAAATGATGATGATGATGATGATGATGATGATGAT	540
541	TTATTAATTTTTCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	600
541	TTATTAATTTTTCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	600
601	GAACATTTAGAAATTTAGTTATGATGATGATGATGATGATGATGATGATGATGATGAT	660
601	GAACATTTAGAAATTTAGTTATGATGATGATGATGATGATGATGATGATGATGATGAT	660
661	TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	720
661	TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	720
721	AATAAAGTACGTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
721	AATAAAGTACGTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780

QY	781	CATAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	840
DB	781	CATAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	840
QY	841	TTATTAGAAATCGTTGATTTA	861
DB	841	TTATTAGAAATCGTTGATTTA	861

RESULT 3

AAS54848

ID AAS54848 standard; DNA; 882 BP.

XX AAS54848;

XX 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #1160.

XX Antisense; ds; prokaryotic cellular proliferation gene;

XX antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU36989.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 8485; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the

XX genes, their use in the discovery of novel antibiotics, the essential

XX genes themselves and the encoded proteins. The prokaryotes used are

XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

XX invention is also useful for the identification of potential new targets

XX for antibiotic development. The antisense nucleic acids can also be used

XX to identify proteins used in proliferation, to express these proteins,

XX and to obtain antibodies capable of binding to the expressed proteins.

XX The proteins can be used to screen compounds in rational drug discovery

XX programmes. The antisense nucleic acid sequence is also useful to screen

XX for homologous nucleic acids which are required for cell proliferation in

XX a wide variety of organisms. The present sequence encodes an

XX essential prokaryotic cellular proliferation protein.

XX Note: the sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic

XX format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 882 BP; 330 A; 116 C; 185 G; 251 T; 0 other;

Query Match 98.3%; Score 846.6; DB 23; Length 882;
Best Local Similarity 99.0%; Pred. No. 9.3e-191;
Matches 852; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGACGAATCTACCGATGAATAATTAATAGATGAAGTCAATATGATTAATCGTTGGC 60
DB 1 ATGACGAATCTACCGATGAATAATTAATAGATGAAGTCAATATGATTAATCGTTGGC 60

QY 61 ATAATAAATCAGTAATGGTACTCAGCTAGAGAAAGTATGTTGTTATTCATTAATGCT 120
DB 61 ATAATAAATCAGTAATGGTACTCAGCTAGAGAAAGTATGTTGTTATTCATTAATGCT 120

QY 121 GGAGTAAACGCATCCAGCAGTCTGTTATTAATCTACTTATTAATCAATACCGAG 180
DB 121 GGAGTAAACGCATCCAGCAGTCTGTTATTAATCTACTTATTAATCAATACCGAG 180

QY 181 TATGAGTTAGTATGAAGAGCGCAATGCTACTAGAAATGATTCATACATATTCACATTAT 240
DB 181 TATGAGTTAGTATGAAGAGCGCAATGCTACTAGAAATGATTCATACATATTCACATTAT 240

QY 241 CATGATGACCTACCAGCGATGGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 CATGATGACCTACCAGCGATGGATGATGATGATGATGATGATGATGATGATGATGAT 300

QY 301 AAGTATATGCTGAGTGGCTGCGATATTAAGCAGGTGATGCTTTATTAATCAATGCTTT 360
DB 301 AAGTATATGCTGAGTGGCTGCGATATTAAGCAGGTGATGCTTTATTAATCAATGCTTT 360

QY 361 GAATTTATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 GAATTTATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

QY 421 CTGCTAATAGCAAGTGGTCTATGTTGGAATGGTGGCGGTGCTGCTGCTGCTGCTGCTG 480
DB 421 CTGCTAATAGCAAGTGGTCTATGTTGGAATGGTGGCGGTGCTGCTGCTGCTGCTGCTG 480

QY 481 GAAGGCCAACCAATGCTTGAACCTTGGAAATGATGATGATGATGATGATGATGATGATG 540
DB 481 GAAGGCCAACCAATGCTTGAACCTTGGAAATGATGATGATGATGATGATGATGATGATG 540

QY 541 TTATTAATCTTTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 CTATTAACATTTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

QY 601 GACATTTAGAAAGTTATAGTTATCAATTTAGTTATGATGTTCCAGATTAAGATGATTTA 660
DB 601 GACATTTAGAAAGTTATAGTTATCAATTTAGTTATGATGTTCCAGATTAAGATGATTTA 660

QY 661 TTAGCTGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
DB 661 TTAGCTGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720

QY 721 AATAAAGTACGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 AATAAAGTACGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 781 CATAGAGACCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CATAGAGACCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

QY 841 TTATTAAGAAATCGTTGATTTA 861
DB 841 TTATTAAGAAATCGTTGATTTA 861

RESULT 4

AAS51602

ID AAS51602 standard; DNA; 864 BP.

XX

AC AAS51602;

XX

DT 13-FEB-2002 (first entry)

XX

DE xx Staphylococcus aureus DNA for cellular proliferation protein #19.
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
OS Staphylococcus aureus.
XX WO200170955-A2.
PN 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US09180.
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
DR P-PSDB; AAU33743.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT Claim 27; Seq ID No 4184; 511pp; English.
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 864 BP; 323 A; 112 C; 182 G; 247 T; 0 other;

Query Match 96.6%; Score 831.6; DB 23; Length 864;
Best Local Similarity 98.9%; Pred. No. 3.3e-187;
Matches 837; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 16 ATGATAAAATTAATAGATGAAGTCAATATGATTAATCGTTGCGGATAAATAATCAGTA 75
DB 1 ATGATAAAATTAATAGATGAAGTCAATATGATTAATCGTTGCGGATAAATAATCAGTA 60

QY 76 ATGGTACTCAGCTAGAGAAAGTATGTTGTTATTAATTAATCGTGAGGTAACGCATC 135
DB 61 ATGGTACTCAGCTAGAGAAAGTATGTTGTTATTAATTAATCGTGAGGTAACGCATC 120

QY 136 CGACCAGTTCTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 195
DB 121 CGACCAGTTCTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180

QY 196 AGACGCCAAATTCGACTAGAAATGATTCATATATTCATTCATGATGACCTACCA 255

Db 191 AAGAGCGCAATGGCACTAGAAATGATTCATACATATTCACCTATTGATGACCTACCA 240
QY 256 CGGATGGATAATGATGATTATCGAGGAGGAAATTAACAAATCATATAAGTATATGGTGAG 315
Db 241 CGGATGGATAATGATGATTATCGAGGAGGAAATTAACAAATCATATAAGTATATGGTGAG 300
QY 316 TGGACTGGATATGAGCAGGTGATGCTTTATTAAGCAATTTGAACCTATTATTCAGT 375
Db 301 TGGACTGGATATGAGCAGGTGATGCTTTATTAAGCAATTTGAACCTATTATTCAGT 360
QY 376 GATGATAGATTAACTGATGAGTAAATAAAGTTCTACACGCTGCTCAATAGCAGT 435
Db 361 GATGATAGATTAACTGATGAGTAAATAAAGTTCTACACGCTGCTCAATAGCAGT 420
QY 436 GGTGATGTTGAATGGTGGCGGTCAATGTTAGATATGCAAGGCGCAACCAAT 495
Db 421 GGTGATGTTGAATGGTGGCGGTCAATGTTAGATATGCAAGGCGCAACCAAT 480
QY 496 GATCTTGAACCTTTGGAATGATACACAAACAAACAGAGCATTTAACTTTTGGC 555
Db 481 GATCTTGAACCTTTGGAATGATACACAAACAAACAGAGCATTTAACTTTTGGC 540
QY 556 GTTATGAGTCAGCAGATATCGTAATGTCGATGATACAACTAAAGACATTTAGAAAGT 615
Db 541 GTTATGAGTCAGCAGATATCGTAATGTCGATGATACAACTAAAGACATTTAGAAAGT 600
QY 616 TATAGTTATCATTTAGTATGATGTTCCAGATTAAGATGATTTATTAAGTCTGCTATGGT 675
Db 601 TATAGTTATCATTTAGTATGATGTTCCAGATTAAGATGATTTATTAAGTCTGCTATGGT 660
QY 676 GATGAAGCAATGATGTTAAAGTGGGCGGAGCATCTTGAATATAAAGTACGTCAC 735
Db 661 GATGAAGCAATGATGTTAAAGTGGGCGGAGCATCTTGAATATAAAGTACGTCAC 720
QY 736 GTGAGTTATTAAGGAAAGATGGCGCAGAGATAAATGACTTATCATAGAGAGCAGCA 795
Db 721 GTGAGTTATTAAGGAAAGATGGCGCAGAGATAAATGACTTATCATAGAGAGCAGCA 780
QY 796 GTGAGTAACTAACGCAATTTGATGAACAATTTCAATACAAACACTTTATTAAGAAATCGTT 855
Db 781 GTGAGTAACTAACGCAATTTGATGAACAATTTCAATACAAACACTTTATTAAGAAATCGTT 840
QY 856 GATTTA 861
Db 841 GATTTA 846

RESULT 5
AV74466/c
AAV74466 standard; DNA; 1893 BP.
AAV74466;
16-MAR-1999 (first entry)
Staphylococcus aureus contig SEQ ID #155.
Computer readable medium; vaccine; S.aureus infection; immunodetection;
cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
skin infection; surgical wound infection; scalded skin syndrome;
toxic shock syndrome; ds.
Staphylococcus aureus.
Key Location/Qualifiers
misc_feature 1261..1320
/tag= a
/note= *these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"

PN EP786519-A2.
PD 30-JUL-1997.
PF 07-JAN-1997; 97EP-0100117.
PR 05-JAN-1996; 96US-0009861.
PX (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX WPI; 1997-374922/35.
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
stored on computer readable medium and used in the production of
anti-S.aureus vaccines
PS Claim 1; Page 784-785; 3271pp; English.
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
of the invention. The DNA sequences are recorded on a computer readable
medium, preferably selected from a floppy or hard disk, random access
memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
the S.aureus DNA sequences allows putative functions to be assigned so
that protein-encoding or regulatory regions of commercial, therapeutic or
industrial importance can be obtained. Specifically, sequences which are
likely to encode antigens have been identified and these polypeptides can
be used in a vaccine composition against S.aureus infection. The
polypeptides can also be used in a kit for the immunodetection of
S.aureus in a sample. S.aureus is implicated in numerous human diseases,
including cellulitis, eyelid infections, food poisoning, osteomyelitis,
skin and surgical wound infections, scalded skin syndrome, toxic shock
syndrome, etc. Organisms transformed with the DNA sequences can be used
for recombinant production of the polypeptides. The new DNA sequences
(and their fragments) are useful as primers or probes for isolating
homologues of any of the S.aureus DNA sequences contained on the
computer readable medium.
SQ Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;
Query Match 93.7%; Score 807; DB 18; Length 1893;
Best Local Similarity 100.0%; Pred. No. 2.5e-181;
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 GTTGGCATAAATAATCAGTAAATGATGATCTAGCTAGAGAAAGTATGTTGATTCATTA 114
Db 1260 GTTGGCATAAATAATCAGTAAATGATGATCTAGCTAGAGAAAGTATGTTGATTCATTA 1201
QY 115 AATGCTGGAGGTAACGATCCGACGATCTCTGTTTATTACTCAGTTTAGATTCACTAAAT 174
Db 1200 AATGCTGGAGGTAACGATCCGACGATCTCTGTTTATTACTCAGTTTAGATTCACTAAAT 1141
QY 175 ACCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 234
Db 1140 ACCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
QY 235 CTTATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 294
Db 1080 CTTATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
QY 295 AATCATAAAGTATATGTTGAGTGGACTCGCATATTTAGCAGGTGATGCTTTTAACTAAA 354
Db 1020 AATCATAAAGTATATGTTGAGTGGACTCGCATATTTAGCAGGTGATGCTTTTAACTAAA 961
QY 355 GCATTGGAACCTATTTCAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 414
Db 960 GCATTGGAACCTATTTCAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
QY 415 CAACGGCTGTCAATAGCAAGTGTCTCATGTTGGAATGGTGGCGGTCAATGTTAGATATG 474
Db 900 CAACGGCTGTCAATAGCAAGTGTCTCATGTTGGAATGGTGGCGGTCAATGTTAGATATG 841

QY 475 CAAAGCGAGGCGCAACCAATTTGATCTTGAACCTTTGGAATGATACACAAAAACAA 534
|||||
Db 840 CAAAGCGAGGCGCAACCAATTTGATCTTGAACCTTTGGAATGATACACAAAAACAA 761
535 GGAGCATTAATACTTTTGGCGTTATGAGTGCAGCAGATATCCCTAATGTCGATATACA 594
|||||
Db 780 GGAGCATTAATACTTTTGGCGTTATGAGTGCAGCAGATATCCCTAATGTCGATATACA 721
595 ACTAAAGAACATTTAGAACTTATAGTATCATTTAGGTATGATGTTCCAGATTAAGAT 654
|||||
Db 720 ACTAAAGAACATTTAGAACTTATAGTATCATTTAGGTATGATGTTCCAGATTAAGAT 661
660 GATTTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
715 GAAATTAATAAGTACGTACGTGATGATGATGATGATGATGATGATGATGATGATGATG 774
600 GAAATTAATAAGTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
775 ACTTATCATAGACGACGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
540 ACTTATCATAGACGACGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
835 AAACACTTATTAGAAATCGTTGATTTA 861
480 AAACACTTATTAGAAATCGTTGATTTA 454

RESULT 6

ABN90871
ID ABN90871 standard; DNA; 909 BP.
AC ABN90871;
XX
XX 24-JUL-2002 (first entry)
XX
XX Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:334.
DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
KW
XX Staphylococcus epidermidis.
OS
XX US6380370-B1.
XX
XX 30-APR-2002.
XX
XX 13-AUG-1998; 98US-0134001.
XX
XX 14-AUG-1997; 97US-055779P.
XX
XX 08-NOV-1997; 97US-064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
XX
XX P-PSDB; ABP38326.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
PT
XX Disclosure; SEQ ID 334; 267pp; English.
XX
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life

CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.

XX Sequence 909 BP; 361 A; 106 C; 165 G; 277 T; 0 other;
Query Match 52.2%; Score 449.8; DB 24; Length 909;
Best Local Similarity 70.2%; Pred. No. 6.2e-97;
Matches 604; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

QY 1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATTAATGAATATCGGTGCG 60
|||||
Db 28 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATTAATGAATATCGGTGCG 87
61 ATAAATAATCACTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
|||||
Db 88 ATAAATAATCACTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 147
121 GGAGGTAATCACTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
148 GGAGGTAATCACTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
181 TATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
208 TATCAACAAGGACTAAATAGTCTTGGATGATGATGATGATGATGATGATGATGATGATG 267
241 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
268 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 327
301 AAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
328 AAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 387
361 GAATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
388 GAATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447
421 CTGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
448 CTTTCAAAAGCAAGTGGACATTTGGGATGATGATGATGATGATGATGATGATGATGATG 507
481 GAAGGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
508 GAAGGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 567
541 TTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
568 TTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 627
601 GAACATTTAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
628 AAGAATTTAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
661 TTAGACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
688 CTGGATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
721 AATAAAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
748 CATAAAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807
781 CATAGAGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
808 CATCAATATCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
841 TTAGAGAAATCGTTGATTTA 861
868 TTAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888

RESULT 7

PT related polypeptides
PS Claim 1; SEQ ID No 1; 192pp; French.
XX The present sequence is the genome sequence of *Listeria monocytogenes*
CC EGD-e. This sequence and fragments of this sequence are useful for
CC selecting probes and primers for detecting genes in *L. monocytogenes* and
CC related organisms, and to study genetic polymorphisms and other genomes.
CC Proteins (AB47297-AB50149) expressed from the present sequence are
CC useful for raising specific antibodies, identification of *L.*
CC *monocytogenes* and related organisms, and for biosynthesis and
CC biodegradation, especially biosynthesis of Vitamin B12. This sequence and
CC proteins encoded by it are also useful for selecting compounds that
CC regulate gene expression and cell replication and modulate *L.*
CC *monocytogenes*-related diseases. In addition, this sequence and proteins
CC encoded by it are useful in pharmaceutical and vaccines compositions for
CC the treatment or prevention of infections by *L. monocytogenes* and related
CC organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;
Query Match 29.3%; Score 252.6; DB 24; Length 2944528;
Best Local Similarity 59.3%; Pred. No. 1.6e-49;
Matches 429; Conservative 0; Mismatches 294; Indels 0; Gaps 0;
QY 61 ATAAATAATCAGTAATGATGATCTCAGCTAGAGAAAGTATGTTTATTTAAATGCT 120
DB 1386000 ATAAAGAGCGGAAATATCGACCTAGACTAAAGAGTCCATGTTATTTCTTCAAGCA 1386059
QY 121 GGAGTAAACGCATCCGACCACTGCTGTTTATTTACTCACTATGATTTCAATTAACCGAG 180
DB 1386060 GCGGGAACGATGCTGCTCAATGCTAGTTTTCGCTACGCTTCAAGCTTTAAAGTAAT 1386119
QY 181 TATGATTTAGTATGAGAGCGCAATTCCTAGTAAATGATTCATATATTCATTTAT 240
DB 1386120 CCGCTTTTATGTTGTAATAACCGCAACGCTTTAGAAATGATTCACATACAGCTTAT 1386179
QY 241 CATGATGATCTACGAGCGATGATTAATGATGATTCACGAGGAAATTAACAATCAT 300
DB 1386180 CATGATGATCTACGAGCGATGATTAATGATGATTCACGAGGCGAGTGCATTAATCAT 1386239
QY 301 AAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 1386240 AAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386299
QY 361 GAACCTATTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 1386300 TCTATTTAGTGAAGACGATTAATTTATCTTTGAGACACGATGCTTTGATTAACCA 1386359
QY 421 CTGTCAATAGCAAGTGGTTCATGTTGGAATGCTGCGCGCTCAATGTTAGATATGCAAGC 480
DB 1386360 ATAGCTTTTATGAGCGTGCAGAGGATGTTGTTGCTCACTGCGAGCTTGAACG 1386419
QY 481 GAAGCCCAACCAATGATCTTTGAACCTTTGGAATGATGATGATGATGATGATGATGATGAT 540
DB 1386420 GAAACAAACAAAGTACGCTAGAGAGTATCATCTTCATTCATGCGAGGAAACGGTGAA 1386479
QY 541 TTATTAACCTTTGCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 1386480 TTATTAATTTATGCTGTAACCTGTCAGCAAAATTTGCGGAGCTGATCCAGAGCAAC 1386539
QY 601 GAACATTTAGAAAGTTATGATTTATCATTTAGTATGATGTTTCCAGATTAAGATGATTTA 660
DB 1386540 AAACCTTACGAAATTTTTCAGAGNATATTTGGGATTTGATTTCAATTTAGCGACCAT 1386599
QY 661 TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 1386600 TTAGATGTAATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386659
QY 721 AATAAAGTACGCTGATGATTTATAGGGAAGATGCGGAGAGATTAATTTGACTTAT 780

DB 1386660 AATAAAGTACCTATCCCGGATTACTACGCTTGATGGGCAAAAAGGCGCAATTAATGAG 1386719
QY 781 CAT 783
DB 1386720 CAT 1386722
RESULT 13
ABQ71010
ID ABQ71010 standard; DNA; 5635 BP.
XX ABQ71010;
XX AC ABQ71010;
XX 29-AUG-2002 (first entry)
XX *Listeria monocytogenes* 4b contig DNA sequence #952.
XX Antibacterial; *Listeria*; food contamination; mutational analysis;
XX infection; ds.
XX *Listeria monocytogenes* 4b.
XX WO200228891-A2.
XX 11-APR-2002.
XX 04-OCT-2001; 2001WO-FR03061.
XX 04-OCT-2000; 2000FR-0012697.
XX (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX Kunst F; Glaser P;
XX WPI; 2002-332479/37.
XX New genomic sequences from *Listeria* species, useful for detection,
XX treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators
XX Claim 14; SEQ ID 3823; 180pp; French.
XX The present invention relates to nucleic acid sequences
XX (ABQ67188-ABQ71212) from *Listeria* sp. The sequences are useful as probes
XX and primers for identification and/or detection of *Listeria* (e.g. as
XX contaminants in foods, or mutational analysis) and for analysis of
XX gene expression. Proteins encoded by the nucleic acid sequences can be
XX used to screen for compounds that modulate gene expression, replication
XX and pathogenicity of *Listeria* (potential therapeutic agents), also for
XX treating infections by *Listeria*, and are useful as immunogens in
XX anti-*Listeria* vaccines.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 5635 BP; 1896 A; 927 C; 1238 G; 1574 T; 0 other;
Query Match 27.9%; Score 239.8; DB 24; Length 5635;
Best Local Similarity 58.2%; Pred. No. 4.2e-47;
Matches 421; Conservative 0; Mismatches 302; Indels 0; Gaps 0;
QY 61 ATAAATAATCAGTAATGATGATCTCAGCTAGAGAAAGTATGTTTATTTCAATTAATGCT 120
DB 2944 ATAAAGAGCGGAAATATCGACCTAGACTAAAGAGTCCATGTTATTTCTTCAAGCA 3003
QY 121 GGAGTAAACGCATCCGACCACTGCTGTTTATTTACTTACTTATGATTTCACTAATTAACCGAG 180
DB 3004 GGTGAAAACGATTCGTCCTCCATGCTAGTTTGTCTACACTTCAAGCCCTTAAAGTAAT 3063
QY 181 TATGATTTAGTATGAGAGCGCAATTCGCTAGAAATGATTCATACATATTCATTTAT 240

Db 3064 CCGCTTTAGTGTGAACAGCAACGGCTTGAAGATGATTCATAGTATGATCCATGATCCGTCATC 3123
Qy 241 CATGATGACCTACCAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 3124 CATGATGATTTACCAGCAATGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3183
Qy 301 AAAGTATATGCTGAGTGGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 3184 AAAGTATATGCTGAGTGGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3243
Qy 361 GAATCTATTTCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 3244 TCATTTTACCGGAACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3303
Qy 421 CTCTCAATAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 3304 ATTAGTATTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3363
Qy 481 GAAGGCCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 3364 GAAACCAACAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3423
Qy 541 TTATTAACCTTTTCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 3424 TTATTAACCTTTTCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3483
Qy 601 GAACATTTTGAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 3484 AAAAGATTTACGGATTTTTCGGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3543
Qy 661 TTAGATGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 3544 TTAGATGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3603
Qy 721 AATAAAGTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 3604 AATAAAGTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3663
Qy 781 CAT 783
Db 3664 CAT 3666

RESULT 14

ABQ67195/c

ID ABQ67195 standard; DNA; 495269 BP.

XX AC

ABQ67195;

29-AUG-2002 (first entry)

Listeria innocua contig DNA sequence #8.

XX DE Listeria innocua contig DNA sequence #8.
XX DE Antibacterial; Listeria; food contamination; mutational analysis;
XX DE Infection; ds.
XX KW

OS Listeria innocua.

XX PN

XX PN WO200228891-A2.

XX PD

XX PD 11-APR-2002.

XX PF 04-OCT-2001; 2001WO-FR03061.

XX PR 04-OCT-2000; 2000FR-0012697.

XX PR (INSP) INST PASTEUR.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

XX XX Kunst F, Glaser P;

XX XX WPI; 2002-332479/37.

XX XX

PT New genomic sequences from *Listeria* species, useful for detection, treatment and prevention of infection, also related polypeptides, antibodies and modulators

PT PT

XX Claim 5; SEQ ID 8; 180pp; French.

XX CC The present invention relates to nucleic acid sequences (ABQ67198-ABQ671212) from *Listeria* sp. The sequences are useful as probes and primers for identification and/or detection of *Listeria* (e.g. as contaminants in foods, or mutational analysis) and for analysis of gene expression. Proteins encoded by the nucleic acid sequences can be used to screen for compounds that modulate gene expression, replication and pathogenicity of *Listeria* (potential therapeutic agents), also for treating infections by *Listeria*, and are useful as immunogens in anti-*Listeria* vaccines.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

CC XX

SQ Sequence 495269 BP; 159529 A; 88849 C; 93135 G; 153754 T; 2 other;

Query Match 26.9%; Score 231.2; DB 24; Length 495269;
Best Local Similarity 56.4%; Pred. No. 1.2e-44;
Matches 431; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 20 ATAAATTAATAGATGAAGTCAATTAATCAATTAATCGGTTCGTAATAAATCAATCAATGATGG 79

Db 366205 ATCATTATATAAAGTACITGATGAGTCGCTTTTAAAGAAATAAATATCGCAATATCG 366146

QY 80 ATACTCAGTAGAAGAAAGTATGTTGATTCATTAATGCTGGAGGTAACCGATCGGAC 139

Db 366145 AACCTAAGTTGAAGAAATCAATTAATTAATCACTGCAAGCTGCGGAAACGAAATTCGTC 366086

QY 140 CAGTCTCTGTTATTAATCACTTTAGATTCATTAATACCGAGTATGATGATGATGAAGA 199

Db 366085 CAATGCTGTTTTCGCAACGCTTCAAGCCTTAATATTAAGCCAAATGCGCGGTTTAAAAA 366026

QY 200 GCGCAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259

Db 366025 CAGCTACGCGCATTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365966

QY 260 TGGATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 319

Db 365965 TGGAT 365906

QY 320 CTGCGATATAGCAGGTGATGCTTTTAACTAAAGCATTGAACTTATTAATCAATGATGATGATGAT 379

Db 365905 CAGCAATTTTAGCAGGAGATGCTTTTAACTAAAGCATTGAACTTATTAATCAATGATGATGATGAT 365846

QY 380 ATAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 439

Db 365845 AAAATCTATCTTTTCGAAACGCGTATTGCTTTAATTAATCAATGATGATGATGATGATGATGAT 365786

QY 440 ATGTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 499

Db 365785 CAGAAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365726

QY 500 TTGAAACTTTTGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559

Db 365725 TAGAAGAACTAGCATCAATCCAGCGCTCGCAAACTGTTGAACTTATTAATCAATGATGATGATGAT 365666

QY 560 TGAGTGCAGCAGATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 619

Db 365665 CCGTCAGCGCAAAATCGCTGAAGCAACTCCAGCAACAAACAAACGATGATGATGATGATGATGAT 365606

QY 620 GTTATCATTTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 679

Db 365605 CAGAAATATCGGCATGTTGTTTCAATTAGCCAGCATATTTTATGATGATGATGATGATGATGAT 365546

QY 680 AAGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739

Db 365545 AAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365486

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 16:10:29 ; Search time 61 Seconds
(without alignments)
4328.667 Million cell updates/sec

Title: US-09-925-637-63

Perfect score: 861
Sequence: 1 atgacgaattaccgatgaa.....tattagaatcgttgattta 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	853	99.1	882	3	US-09-276-873-1
2	449.8	52.2	909	4	US-09-134-001C-334
3	157.2	18.3	894	1	US-08-534-910B-5
4	155.6	18.1	893	1	US-08-534-910B-3
5	155.6	18.1	894	1	US-08-534-910B-2
6	155.6	18.1	894	1	US-08-534-910B-3
7	155.6	18.1	894	1	US-08-534-910B-4
8	155.6	18.1	894	3	US-08-886-466-1
9	155.6	18.1	894	4	US-09-475-304-1
10	155.6	18.1	894	4	US-09-101-126-4
11	155.6	18.1	894	4	US-09-367-528A-4
12	154	17.9	894	1	US-08-534-910B-1
13	154	17.9	894	4	US-09-367-528A-2
14	150.6	17.5	486	3	US-09-217-609A-28
15	150.6	17.5	486	4	US-08-873-235B-28
16	136.4	15.8	885	3	US-09-187-050-11
17	134.8	15.7	1179	3	US-09-187-050-13
18	134.8	15.7	1179	3	US-09-187-050-15
19	134.8	15.7	1179	3	US-09-187-050-17
20	134.8	15.7	1179	3	US-09-187-050-19
21	134.8	15.7	1179	3	US-09-187-050-21
22	134.8	15.7	1179	3	US-09-187-050-23
23	134.8	15.7	1179	3	US-09-187-050-25
24	134.8	15.7	1889	3	US-09-187-050-1
25	124.6	14.5	876	3	US-09-275-742-1
26	124.2	14.4	10011	4	US-08-961-527-76
27	76.8	8.9	1284	4	US-08-858-207A-173

28	72.6	8.4	861	3	US-09-187-050-4	Sequence 4, Appli
29	72	8.4	909	1	US-07-783-705A-7	Sequence 7, Appli
30	72	8.4	6918	1	US-07-783-705A-13	Sequence 13, Appli
31	71.4	8.3	990	1	US-08-410-167A-1	Sequence 1, Appli
32	71.4	8.3	993	1	US-08-705-377-1	Sequence 1, Appli
33	71.4	8.3	993	1	US-08-705-377-2	Sequence 2, Appli
34	71.4	8.3	993	1	US-08-705-377-4	Sequence 4, Appli
35	71.4	8.3	993	1	US-08-705-377-5	Sequence 5, Appli
36	71.4	8.3	993	2	US-09-052-962-1	Sequence 1, Appli
37	71.4	8.3	993	2	US-09-052-962-2	Sequence 2, Appli
38	71.4	8.3	993	2	US-09-052-962-4	Sequence 4, Appli
39	71.4	8.3	993	2	US-09-052-962-5	Sequence 5, Appli
40	71.4	8.3	993	2	US-09-053-068-1	Sequence 1, Appli
41	71.4	8.3	993	2	US-09-053-068-2	Sequence 2, Appli
42	71.4	8.3	993	2	US-09-053-068-4	Sequence 4, Appli
43	71.4	8.3	993	2	US-09-053-068-5	Sequence 5, Appli
44	71.4	8.3	993	2	US-08-898-560-2	Sequence 2, Appli
45	71.4	8.3	993	4	US-09-101-126-2	Sequence 2, Appli

RESULT 1
US-09-276-873-1
; Sequence 1, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: Ispa
; FILE REFERENCE: GM10208
; CURRENT APPLICATION NUMBER: US/09/276.873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-276-873-1

Query Match 99.1%; Score 853; DB 3; Length 882;
Best Local Similarity 99.4%; Pred. No. 2.6e-198;
Matches 856; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy	1	ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTAATCGGTGCG	60
Db	1	ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTAATCGGTGCG	60
Qy	61	ATAAATAAATCAGTAATGGATCTCAGCTAGAGAAAAGATGTGTATTCTATTAAATGCT	120
Db	61	ATAAATAAATCAGTAATGGATCTCAGCTAGAGAAAAGATGTGTATTCTATTAAATGCT	120
Qy	121	GGAGTAAACCCATCCGACCGAGTTCTGTATTACTTACTTACTTACTTACTTACTTACTT	180
Db	121	GGAGTAAACCCATCCGACCGAGTTCTGTATTACTTACTTACTTACTTACTTACTTACTT	180
Qy	181	TATCAGTATAGTATGAAGAGCGCAATTCCTAGTAAATGATTCATACATATTCATTATT	240
Db	181	TATCAGTATAGTATGAAGAGCGCAATTCCTAGTAAATGATTCATACATATTCATTATT	240
Qy	241	CATGATGACCTTACCAGCGATGAATAATGATTCATGACGAGGAAAATTAACAAATCAT	300
Db	241	CATGATGACCTTACCAGCGATGAATAATGATTCATGACGAGGAAAATTAACAAATCAT	300
Qy	301	AAAGTATATGAGTGGAGTGGCGATATAGCAGGTGATGCTTTATTAACTAAACGATTT	360
Db	301	AAAGTATATGAGTGGAGTGGCGATATAGCAGGTGATGCTTTATTAACTAAACGATTT	360
Qy	361	GAAGTATTTCAAGTATGATAGATTAACTGATGAGTAAATAAATAAAGTTCTACACGG	420
Db	361	GAAGTATTTCAAGTATGATAGATTAACTGATGAGTAAATAAATAAAGTTCTACACGG	420

ALIGNMENTS


```
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: IBM/Word Perfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/534,910B
FILING DATE: 28-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-25253
FILING DATE: 14-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 77670/398
TELEPHONE: (202)429-1776
TELEFAX: (202)429-0796
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
ORIGINAL SOURCE:
ORGANISM: Bacillus stearothermophilus
US-08-534-910B-5

Query Match      18.3%; Score 157.2; DB 1; Length 894;
Best Local Similarity 51.0%; Pred. No. 1.5e-29;
Matches 400; Conservative 0; Mismatches 378; Indels 6; Gaps 1;

QY 17 TGAATAAATTAATAGATGAAGTCAATATGAATATATCGGTTCGCGATAAATAATCAGTAA 76
DB 29 TCAACGAGCAAAACAGCGGTGGAACACGCGCTCCCGTTATATAGAGCGCTTAGAAG 88
QY 77 TGGATACCTACGTAGCAAGAGTATGTTGTTATTAATTAATGCTGGAGGTAAACGATCC 136
DB 89 GCCCGGCAAGCTGAAAGGCGATGCGGTACTATTGGAGCGCGCGGCAACGAAATCC 148
QY 137 GACCAGTTCCTGTTACTACTTATAGATTCACTAAATPACCGAGTATGAGTTAGTATGA 196
DB 149 GTCCGTTGCTGCTCTGCTCCACCGTTCAGCGCTCGGCAAGACCCGCGTCGGATTGC 208
QY 197 AGAGCGCAATTCGACTGAAGATGATCATATATCTACTTATCATGATGACCTACACG 256
DB 209 CCGTCGCTGCGCGGATGAAATGATCCATAGCTACTCTTTGATCCATGATGATTTGCCGA 268
QY 257 CGATGGCAATGATGATTATCGACGAGGAGAAATTAACAAATCAATAAGTATATGTTGAGT 316
DB 269 GCATGGCAACGATGATTGCGGCGCGGCAAGCGGCAACCACTAAGTGTTCGCGGAGG 328
QY 317 GGAATCGCATATAGCAGGTGATGCTTTTAACTAAAGCATTTTGAACCTTATTTTC----- 371
DB 329 CGATGGCCATCTTGGCGGGGACGGTGTGTTGACGTACGCGCTTCAATTTGATCAGCGAAA 388
QY 372 -AAGTATGATAGATTAATCACTGATGAAGTAAATAAAGTTCTACACGCGTGTCAATAG 430
DB 389 TCGACGATGACGCGATCCCTCTCCGTCGCGGCTTCGCGCTCATCGAACGCGTGGCGAAG 448
QY 431 CAAGTGGTCATGTTGGAATGTCGCGGTCAAAATGTTAGATATGCAAGCGAAGGCCAAC 490
DB 449 CGGCGGTTCGGAAGGATGTCGCGGTAGGCGGATGATGAAGGAGAGGGGAAA 508
QY 491 CAATTTGATCTTGAACCTTTGAAATGATACACAAACAAACAGGAGCATTTATTAACTT 550
DB 509 CGCTGACGCTTTCGAGGCTCGAATACATTCATCGGCATAAACCAGGAAATGCTGCAAT 568
QY 551 TTGCGGTTATGAGTGCAGCATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
DB 569 ACAGCGTGCACGCGCGGCGCTTGGATCGCGCGCGCTGATGTCGCGGCAACGCGGGAGCTTG 628
QY 611 AAGTTATAGTTATCAATTTAGGTATGATGTTCCAGATTAAGATTAATTTATAGACTGCT 670
```

```
Db 629 ACGAATTCGCGCCCATCTAGCCCTTGCTTTCAAATTCGCGATGATATCTCGATATG 688
QY 671 ATGGTGATGAAGCAAAAGTTAGTAAAGTGGCGAGCATCTTGAATAATAATAAGTA 730
DB 689 AAGGGCAGAGAAATCGCAAGCGGTGGCGACCAACACAAACGCA 748
QY 731 CGTACGTGAGTTATTAGGAAAGATGGCGCAGAGATAAATTAATTAATCATAGAGAG 790
DB 749 CGTATCCAGCGTTGCTGCTTGGCGCGGCAAGGAAAGTTGGGTTCCATCATGAGG 808
QY 791 CAGC 794
DB 809 CGC 812

RESULT 4
US-08-333-321-1
; Sequence 1, Application US/08333321
; Patent No. 5786192
; GENERAL INFORMATION:
; APPLICANT: Odata, Shusei
; TITLE OF INVENTION: FARNESYL PYROPHOSPHATE SYNTHETASE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward W. Greason, Esq.
; STREET: 1 Broadway
; CITY: New York
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: No. 5786192epad, Windows 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,321
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/953,424
; FILING DATE: 29-SEP-1992
; APPLICATION NUMBER: JP 3-253788
; FILING DATE: 01-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Greason, Edward W.
; REGISTRATION NUMBER: 18,918
; REFERENCE/DOCKET NUMBER: 077670/00310
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 893 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
US-08-333-321-1

Query Match      18.1%; Score 155.6; DB 1; Length 893;
Best Local Similarity 50.9%; Pred. No. 3.7e-29;
Matches 399; Conservative 0; Mismatches 379; Indels 6; Gaps 1;

QY 17 TGAATAAATTAATAGATGAAGTCAATATGAATTAATTCGTTCCGATAAATAATCAGTAA 76
DB 29 TCAACGAGCAAAACAGCGGTGGAACACGCGCTCTCCGTTATATAGAGCGCTTAGAAG 88
QY 77 TGGATACCTACGTAGCAAGAGTATGTTGTTATTAATTAATGCTGGAGTAAACGATCC 136
DB 89 GCCCGGCAAGCTGAAAGGCGATGCGGTACTATTGGAGCGCGGCAACGAAATCC 148
```

QY	137	GACCA	TTCTCTG	TATTACT	CACTT	CTAGAT	TCACT	TAATAC	CGAG	TATG	AGTTAG	GTATGA	196
Db	149	GTCCG	TGCTG	CTTCTG	TCTCA	CCGTT	CGGCG	TCTCG	CAAA	GACCG	GGGGT	CGGATTGC	208
QY	197	AGAGC	GAATTC	CACTAG	AAATG	ATTC	ACAT	ATTT	CAC	TATTC	ATGAT	GACCTACC	256
Db	209	CCGTC	CGCTCG	CGGAT	TGAAT	GAATC	CACT	ACGT	ATCT	TTGAT	TCCA	TGATTTGCCG	268
QY	257	CGATG	GAATAT	CATG	ATTAT	TCG	AGAG	AAATTA	ACA	ATCAT	ATAAG	TATATG	316
Db	269	GCATG	SACA	CGATG	ATTG	CGGCG	CGCA	AGCCG	AGCA	ACCATA	AAAGT	TTTCGG	328
QY	317	GGACT	GC	GATAT	TAG	AGGT	GATG	CTTT	TATTA	CTAA	AGCAT	TTGA	371
Db	329	CGATG	GC	CACT	TTTGG	CGGGG	GAC	GGTGT	TGAC	GTAC	CGGTT	CAAT	388
QY	372	-AAGT	GATG	ATAG	ATTA	ACTG	ATCA	AGTAA	AAAT	TAAAG	TTCT	ACA	430
Db	389	TCGAC	GATG	AGCG	ATCC	CTCCT	CCGTC	CGGCT	TCG	GGCTC	ATCG	AAAGCG	448
QY	431	CAAGT	GGTCA	TGTTG	GAAT	GGTC	GGGT	CAAA	TGTT	AGATAT	GC	AAAGCG	490
Db	449	CGCCG	GTCC	GAGG	GGAT	GTG	CCCGT	CAGG	CCG	GATATG	GAAG	AGGGGA	508
QY	491	CAATT	GTAT	CTTGA	AACT	TTTGA	AAATG	ATAC	AAAA	CAAAA	CAGAG	CAATTA	550
Db	509	CGCT	CAC	GC	TTTTC	GGAG	CTCG	AAAT	TTCAT	GGCAT	AAAA	CCGGAAAA	568
QY	551	TTGCG	GTAT	GAGT	GC	AG	CAGAT	ATCG	CTAAT	TCGAT	GATACA	CTA	610
Db	569	ACAGC	GTGC	AC	CGCG	CCCT	TGAT	CGGCG	CGCT	GATGCC	GGCAAC	CGGGAG	628
QY	611	AAAGT	TATAG	TATCA	TTTAG	GTAT	GTAT	TCC	AGAT	TAAAG	ATGATTT	TAT	670
Db	629	ACGAAT	TCG	CGCC	CACT	AG	CGCT	TG	CCCTT	CAAA	TTCCG	GATATTC	688
QY	671	ATG	TGAT	GAGCA	AAAGT	TAGG	TAAAA	AGTGG	GCG	CGCAT	CTTGA	AAATATA	730
Db	689	AAGGG	CGAG	AGAAAA	ATCG	CAAG	CCGTC	GGC	ACG	CAACCA	CAACA	AGCGA	748
QY	731	CGTAC	GTGAG	TTTAT	TAGG	GAAG	ATG	CGC	AGA	GATAA	ATTTG	ACT	

RESULT 5
US-08-534-910B-2
; Sequence 2, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable
; TITLE OF INVENTION: Of Synthesizing Geranylgeranylidiphosphate And Gene Coding Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.
; ZIP: 20036-5405
; COMPUTER READABLE FORM:

QY 611 AAGCTTATAGTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
DB 629 ACGAATTCGGCGCCCTAGGCTTTCCTTCAATTCGGATGATATTCGATATTG 688
QY 671 ATGCTGATGAACAACTAGTAAAGTGGGCGAGGATCTTGAATAATAAAGTA 730
DB 689 AAGGGCAGAGAAATAATCGGCAAGCGGTCGGCAGCAGCAACCAAGCAACAAAGCGA 748
QY 731 CGTACGTGAGTTTATTAGGAAAGATGGCGCAGAGATAAATGACTTATCATAGAGCG 790
DB 749 CGTATCCAGCGTTGCTGCTGCTTCCGCGCGAAGGAAAGTTGGCGTTCCATATCGAGG 808
QY 791 CAGC 794
DB 809 CGGC 812

RESULT 6

US-08-534-910B-3
Sequence 3, Application US/08534910B
Patent No. 5766911

GENERAL INFORMATION:

APPLICANT: KOIKE, Ayumi
APPLICANT: OBATA, Shusei
APPLICANT: NISHINO, Tokuzo
APPLICANT: OHNUMA, Shinichi
APPLICANT: NAKAZAWA, Takeshi
APPLICANT: OGURA, Kyoza
APPLICANT: KOYAMA, Tanetoshi

TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable

TITLE OF INVENTION: Of Synthesizing Geranylgeranylidiphosphate And Gene Coding Thereof

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: 1025 Connecticut Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: U.S.

ZIP: 20036-5405

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.25" Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS 6.2

SOFTWARE: IBM/Word Perfect 6.1 Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/534.910B

FILING DATE: 28-SEPT-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-25253

FILING DATE: 14-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Toffenetti, Judith L.

REGISTRATION NUMBER: 39,048

REFERENCE/DOCKET NUMBER: 77670/398

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)429-1776

TELEFAX: (202)429-0796

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 894 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA

ORGANISM: Bacillus stearothermophilus

US-08-534-910B-3

Query Match

Best Local Similarity 18.1%; Score 155.6; DB 1; Length 894;

Matches 399; Conservative 0; Mismatches 379; Indels 6; Gaps 1;

QY 17 TGAATAAATTAATAGATGAAGTCAATTAATGAATTAATCGGTTCGATAAATAAATCAAGTAA 76
DB 29 TCAACAGGAGCAAAAACAGCGCTGGAAACAGCGCTTCCCGTTATATAGAGCGCTTAGAAG 88
QY 77 TGGATACCTCAGTAGAAGAAAGTATGTTGTTATTTCAATTAATGCTCGAGGTAAACCATCC 136
DB 89 GCGCGGCAAGCTGAAAAGGCGATGGCGTACTCAATTTGGAGCGCGCGCAACAAATCC 148
QY 137 GACCAAGTTCTTATTACTCAGTTTAGATTTCCTAAATACCGAGTATGAGTTAGTATGA 196
DB 149 GTCCGTGCTGCTTCTGTCACCGTTCCGGCGCTCGGAAAGACCCGCGGTCGGATTC 208
QY 197 AGAGCGCAATTCAGTCAATTAATGATATATATTTCAATTTCAATTTCAATTTCAATTT 256
DB 209 CCGTCGCTGCGCGATTTGAAATGATCCTACTCTTTGATCCTATGATTTGCGGA 268
QY 257 CGATGGATAATGATGATTATCGACGAGGAAATTAACAAATCAATCAATCAATCAATCAAT 316
DB 269 GCATGGACAACGATTTTGGCGCGGCAAGCCGCAACCAATCAATCAATCAATCAATCAAT 328
QY 317 GGAATGCGATATTAGCAGGTGATGCTTTTAACTAAAGCAATTTGAATTTTGAATTTTGA 371
DB 329 CGATGGCCATCTTGGCGGGGACGGTGTGACGCTACGCGTTTCAATTTGATCAGCGAA 388
QY 372 -AAGTATGATAGATTAATTAATGATGAAGTAAATAAATAAAGTTCTACAACGGCTGTC 430
DB 389 TCGACGATGAGCGCATCCCTCTTCCGCTCGGCTTCGCTCATCGAAGCGCTGGCGAAG 448
QY 431 CAAGTGGTCATGTTGGAATGTCGCGCTCAATGTTAGATATGCAAGCGAGGCGCCAC 490
DB 449 CGCGCGTCCGGAAGGATGGCGCGGTCAGCGAGCGGATGGAAGGAGAGGGGAA 508
QY 491 CAATTTGATCTTGAATTTGGAATGATACACAAACAAACAGGAGCATTTAATTA 550
DB 509 CGCTGACGCTTCGGAGCTCGAATACATTCGCTATAAACCAGGAAATGCTGCAAT 568
QY 551 TTGCGGTTATGATGTCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
DB 569 ACAGCGTCACGCGCGGCTTGTGCGGCGGCTGTGCGCGGCAACCGCGGAGCTTG 628
QY 611 AAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAGATGATTTATTAGACTGCT 670
DB 629 ACGAATTCGCGCCCATCTAGGCTTGGCTTTCAATTTCCGATGATATTTCTCGATATTG 688
QY 671 ATGCTGATGAAGCAAGTTAGTTAAAGTGGCGAGCGATCTTGAATAATAAAGTA 730
DB 689 AAGGGCAGAGAAATAATCGCAAGCGGTCGGCAGCGCAACCAAGCAACAAAGCGA 748
QY 731 CGTACGTGAGTTTATTAGGAAAGATGGCGCAGAGATAAATGACTTATCATAGAGCG 790
DB 749 CGTATCCAGCGTTGCTGCTGCTTGGCGCGCAAGGAAAGTTGGCGTTCCATATCGAGG 808
QY 791 CAGC 794
DB 809 CGGC 812

RESULT 7

US-08-534-910B-4

Sequence 4, Application US/08534910B

Patent No. 5766911

GENERAL INFORMATION:

APPLICANT: KOIKE, Ayumi

APPLICANT: OBATA, Shusei

APPLICANT: NISHINO, Tokuzo

APPLICANT: OHNUMA, Shinichi

APPLICANT: NAKAZAWA, Takeshi

APPLICANT: OGURA, Kyoza

APPLICANT: KOYAMA, Tanetoshi

TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable

TITLE OF INVENTION: Of Synthesizing Geranylgeranylidiphosphate And Gene Coding Thereof

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

Query Match	18.18;	Score 155.6;	DB 3;	Length 894;
Best Local Similarity	50.9%;	Pred. No. 3.7e-29;		
Matches 399;	Conservative 0;	Mismatches 379;	Indels 6;	Gaps 11
Qy	17	TGAATAAATTAATAGATGAAGTCAATTAATGAATTAATCGGTTGCAGTAATAATAACAGTAA	76	
Db	29	TCAACGAGCAAAAACACGCGGTGAAACACAGCGCTCTCCGTTATATAGAGCGCTTAGAAG	88	
Qy	77	TGATACTACCTAGCTAGAAAGATGTGTTGATTCATTAATGCTGGAGGTTAAACGCATCC	136	
Db	89	GCGCGCGGAAGCTGAAAAGCGGATGCGGTACTCATTTGGAGCGCGCGGCAAGCAGATCC	148	
Qy	137	GACCAGTTCCTGTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGTATGA	196	
Db	149	GTCCGTTGCTGCTTCTGTCCACCGTTCCGGCGCTCGGCAAGACCCGCGCTCGGATGC	208	
Qy	197	AGAGCGCAATTGCACTAGAATGATTCATACATATTCATCTATTTCATGATGACCTACCAG	256	
Db	209	CCGTGCGCTCGCGGATTTGNAATGATCATACGTACTCTTTGATCCATGATGATTTGCCGA	268	
Qy	257	CGATGGATAATGATGATTATTCGACGAGGAGAAAATTAACAAAATCATATAAGTATATGCTAGT	316	

Db 269 GCATGGACACGATGATTTGGCGCGCGCAAGCCGACGACCACTAAAGTGTTCGGCGAGG 328
QY 317 GGACTCGCATATTAGCAGGTGATCTTTTAACTAAAGCATTTGAACCTTATTTC----- 371
Db 329 CGATGGCCATCTTGGCGGGGACGGGTGTTGACGTACGCTTCAATTGATCACCGAAA 388
QY 372 -AAGTATGATGATTAAGTATGATGAAGTAAATATAAAGTCTTACACGCGCTGTCAATAG 430
Db 389 TCGACGATGACGCATCCCTCTTCCGTCGGGCTTCGGCTCATCGAACGCTGGCGAAAG 448
QY 431 CAAGTGCTCATGTTGGAATGTCGGCGGTCAAAATGTTAGATATGCAAGCGAAGGCCAAC 490
Db 449 CGCGCGTCCGAAGGATGCTCGCGTCAGCGACCGCATATGGAAGGAGAGGGGAAA 508
QY 491 CAATTGATCTTTGAACCTTTGGAATGATACACAAAAACAAAGAGGAGCATTTTAACTT 550
Db 509 CGCTGACGCTTCGGAGCTCGATACATTCATCGGCATAAACCGGGAATGCTGCAT 568
QY 551 TTGCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610
Db 569 ACAGCGTGCACGCGCGCTTTCGCGCGCGCTGATGCGCGCGCTGATGCGCGCGCTGAT 628
QY 611 AAAGTTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
Db 629 ACGAATTCGCGCGCTTTCGCGCGCTTTCGCGCGCTTTCGCGCGCTTTCGCGCGCTTTC 688
QY 671 ATGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
Db 689 AAGGGCAGAGAAATAATCGGACGCGCTGCGACGCGACCAAGCAACACACAGCGA 748
QY 731 CGTACGTGATTTATGAGGAAGATGCGCGAGAGATGATGATGATGATGATGATGATGATGAT 790
Db 749 CGTATCCAGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808
QY 791 CAGC 794
Db 809 CGGC 812

RESULT 9

US-09-475-304-1
; Sequence 1, Application US/09475304
; Patent No. 6225096
; GENERAL INFORMATION:
; APPLICANT: Narita, Keishi
; APPLICANT: Narita, Chika
; APPLICANT: Takeuchi, Yoshie
; APPLICANT: Ohto, Chikara
; APPLICANT: Ohnuma, Shinichi
; APPLICANT: Nishino, Tokuzo
; TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
; FILE REFERENCE: 77670/494
; CURRENT APPLICATION NUMBER: US/09/475,304
; CURRENT FILING DATE: 1999-12-30
; EARLIER APPLICATION NUMBER: JP 8-191635
; EARLIER FILING DATE: 1996-07-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(891)
US-09-475-304-1
Query Match 18.1%; Score 155.6; DB 4; Length 894;
Best Local Similarity 50.9%; Pred. No. 3.7e-29;
Matches 399; Conservative 0; Mismatches 379; Indels 6; Gaps 1;
QY 17 TGATAAATAATAGATGAAGTCAATTAATGAATATCGGTGCGATATAAATAACAGTAA 76

Db 29 TCACGAGCAAAAACAGCGGTGGAACACAGCGCTCTCCGTTATATAGAGCGCTTAGAAG 88
QY 77 TGGATATCTAGCTAGAAGAAAGTATGTTGATTCATTAAATGCTGGAGGTAAACGCAATCC 136
Db 89 GGCGCGCAAGCTGAAAAGCGGATGCGGTACTCATTTGGAGCGCGCGCAACGATATCC 148
QY 137 GACCAAGTTCGTTTATTAATCTACATTTAGTATTCACATAACGAGATGATGATGATGATGAT 196
Db 149 GTCCGTTGCTGCTTCTGCTCCACCTTCGGGCGCTTCGGCAAGACCCGCGGTCGATTCG 208
QY 197 AGACGCGCAATTCGACTAGAAATGATTCATACATATTCATTTATTCATGATGATGATGATGAT 256
Db 209 CCGTCCGCTCGCGGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 268
QY 257 CGATGGATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 316
Db 269 GCATGGACACGATGATTTGGCGCGCGCAAGCGCGCAACCATTAAGTGTTCGCGGAGG 328
QY 317 GGACTCGGATTTAGCAGGTGATGCTTTTAACTAAAGCATTTGAACCTTATTTC----- 371
Db 329 CGATGGCCATCTTTCGCGGCGCGGCTGTTGACGTACGCGTTCATTTCAATTTGATCACCGAAA 388
QY 372 -AAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 430
Db 389 TCGACGATGACGCGCATCCCTCTTCCTCCGCGCTTCGCGCTCATCGAACGCGTGGCGAAG 448
QY 431 CAAGTGCTCATGTTGGAATGTCGGCGGTCAAAATGTTAGATATGCAAGCGAAGGCCAAC 490
Db 449 CGCGCGTCCGAAGGATGTCGCCGTCAGCGACGCGCATATGGAAGGAGAGGGGAAA 508
QY 491 CAATTGATCTTTGAACCTTTGGAATGATACACAAAAACAAAGAGGAGCATTTTAACTT 550
Db 509 CGCTGACGCTTTCGCGGCTCGAATACATTCATCGGCATAAACCGGGAATGCTGCAAT 568
QY 551 TTGCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610
Db 569 ACAGCGTGCAGCGCGCGCTTTCGCGCGCGCTGATGCGCGCGCTGATGCGCGCGCGCGGAGCTTG 628
QY 611 AAAGTTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
Db 629 ACGAATTCGCGCGCTTTCGCGCGCTTTCGCGCGCTTTCGCGCGCTTTCGCGCGCTTTC 688
QY 671 ATGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
Db 689 AAGGGCAGAGAAATAATCGGACGCGCTGCGACGCGACCAAGCAACACAAAGCGA 748
QY 731 CGTACGTGATTTTATGAGGAAGATGCGCGAGAGATGATGATGATGATGATGATGATGATGAT 790
Db 749 CGTATCCAGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808
QY 791 CAGC 794
Db 809 CGGC 812

RESULT 10

US-09-101-126-4
; Sequence 4, Application US/09101126
; Patent No. 6316216
; GENERAL INFORMATION:
; APPLICANT: OHTO, CHIARA
; APPLICANT: NAKANE, HIROYUKI
; APPLICANT: NISHINO, TOKUZO
; APPLICANT: OHNUMA, SHINICHI
; APPLICANT: HIROOKA, KAZUTAKE
; TITLE OF INVENTION: MUTATED PRENYL DIPHOSPHATE SYNTHASES
; FILE REFERENCE: 77670/566
; CURRENT APPLICATION NUMBER: US/09/101,126
; CURRENT FILING DATE: 1999-04-27
; EARLIER APPLICATION NUMBER: PCT/JP97/03921
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: JP 8-307506

QY 731 CGTACGTCAGTTTATTAGGAAAGATGCCGCGAGAAAGATAAATTCAGTCTATCATAGAGAG 790
|||||
Db 749 CGTATCCAGCGTTGCTGTCTGCTTCCGCGCGGAAAGAAAGTTGCGGTTCCATATCGAGG 808

QY 791 CAGC 794
|||
Db 809 CGGC 812

RESULT 12

US-08-534-910B-1
; Sequence 1, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable
; TITLE OF INVENTION: Of Synthesizing Geranylgeranylidiphosphate And Gene Coding Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: IBM/Word Perfect 6.1 Windows
; CURRENT APPLICATION DATA: US/08/534,910B
; FILING DATE: 28-SEPT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-25253
; FILING DATE: 14-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Toffenetti, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 77670/398
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)429-1776
; TELEFAX: (202)429-0796
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
US-08-534-910B-1

Query Match 17.9%; Score 154; DB 1; Length 894;
Best Local Similarity 50.8%; Pred. No. 8.9e-29;
Matches 398; Conservative 0; Mismatches 380; Indels 6; Gaps 1;

QY 17 TGAATAATTAATAGATCACTCAATTAATGATTCGGTTGCCGATTAATAATCAGTAA 76
|||||
Db 29 TCAACGAGCAAAACAGCGGTGGAACAGCGCTCTCCGGTTATATAGAGCGCTTAGAAG 88

QY 77 TGGATCTACGCTAGAGAAAGTATGTTGATTCATTAAATGCTGGAGGTAACGCAATCC 136
|||||
Db 89 GGCCGGCGAAGCTGAAAAGCGGATGCGTACTCATTTGGAGCGCGCGGCAACGATCC 148

QY 137 GACCAGTTCTGTTTACTCTACTTTAGATTCACTAAATACCAGATATCAGTTAGTATGA 196
|||||
Db 149 GTCGGTTGCTGCTTCTGTCACCGTTCCGGGCGCTCGGAAAGACCCCGGTCGGATTGC 208

QY 197 AGAGCGCAATTCGACTAGAAATGATTCATACATATTCATTATTCATGATGACCTACCAG 256
|||||
Db 209 CCGTCGCTCGCGGATGAAATGATTCATACGCACTCTTTGATCCATGATGATTTGCCGA 268

QY 257 CGATGGATAATGATGATTATCGACGAGGAAATTAACAAATCATAAAGTATATGTTGAGT 316
|||||
Db 269 GCATGGACAACGATGATTTGCGCGCGGCAAGCCACCAACCAATAAAGTTTCGCGGAGG 328

QY 317 GGACTGCCATATAGCAGGTGATGCTTTTATTAACATAAAGCAATTTGAATTTATTC- 371
|||||
Db 329 CGATGGCCATCTTGGCGGGGACGGGTTGTTGACGTACGCGTTTCAATTTGATCACCAGAA 388

QY 372 -AAGTGATGATAGATTAACATGATGAAGTAAATAAATAAAGTTCTACAAAGCGCTGTCATAG 430
|||||
Db 389 TCGACGATGAGCGCATCCCTCTCCGTCGCGCTTCGCGCTCATCGAACGCTGGCGAAAG 448

QY 431 CAAGTGGTCATGTTGGAATGGTCGCGGCTCAAAATGTTAGATATGCAAAAGCGGAGCCAC 490
|||||
Db 449 CGCGCGTCCGGAAGGATGTCGCGCTCAGGCAGCGGATATGGAAGAGAGGAGGAA 508

QY 491 CAATTGATCTTTGAAACTTTGGAATGATACACAAACAAACAGAGACATTTAACTT 550
|||||
Db 509 CGCTGACGCTTTCCGAGCTCGAATATCATTCGTCATATAAACCGGAAATGCTGCAAT 568

QY 551 TTGCGTTATGATGTCACGATATCGCTTAATGTCGATGATACACTAAAGAACATTTAG 610
|||||
Db 569 ACAGCGTCACGCGCGCTTGTGTCGCGGCTGATGCGCGGCAACCGCGGAGCTG 628

QY 611 AAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAAGATGATTTATAGACTGCT 670
|||||
Db 629 ACGAATTCGCGCCCATCTAGCGCTTGCCTTCAATTCGCGATGATTTCTCGATATG 688

QY 671 ATGGTGATGAAGCAAAAGTTAGTATAAAGAGTGGGAGCGATCTTGAAATATAAAGTA 730
|||||
Db 689 AAGGGCAGAAAGAAATAATCGCAAGCGGCTCGGAGCGGCAACCAACCAAGCGA 748

QY 731 CGTACGTCAGTTTATTAGGAAAGATGCGCGAGAGATAAATGACTTATCATAGAGAG 790
|||||
Db 749 CGTATCCAGCGTTGCTGTCGCTTCCGCGCGGAGGAAAGTTGCGGTTCCATATCGAGG 808

QY 791 CAGC 794
|||
Db 809 CGGC 812

RESULT 13

US-09-367-528A-2
; Sequence 2, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
; APPLICANT: TOYOTA JIDOSHA KABUSHIKI KAISHA
; TITLE OF INVENTION: Geranyl Diphosphate Synthetase Gene
; FILE REFERENCE: PH-586
; CURRENT APPLICATION NUMBER: US/09/367,528A
; CURRENT FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: JP97/346686
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 894-
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(894)
US-09-367-528A-2

Query Match 17.9%; Score 154; DB 4; Length 894;
Best Local Similarity 50.8%; Pred. No. 8.9e-29;
Matches 398; Conservative 0; Mismatches 380; Indels 6; Gaps 1;

QY 17 TGAATAAATTAATAGATGAAGTCAATTAATATATCGTTGGATATAAATACAGTAA 76
DB 29 TCAACGAGCAAAACAGCGGTGGAACACGCGTCTCCCGTTATATAGAGCGGTAGAG 88

QY 77 TGGATACCTACGCTAGAGAAAGTATGTTGTTATTCATTAAATGCTGGAGTAAACGCGATCC 136
DB 89 GCGCGGGAAGCTGAAAAGCGCATGCGTACTCATTTGGAGCGCGCGCAACAGATCC 148

QY 137 GACAGTTCGTTTATTAATCACTTTAGATTCATTAATACCGAGTATGATGATGATGA 196
DB 149 GTCCGTTGCTGCTCTGTCACCGCTTCGGCGCTCGCAAGACCGCGGCTCGGATTC 208

QY 197 AGAGCGCAATTCGACTAGAAATGATTCATACATATTCACCTTATTCATGATGACCTACCAG 256
DB 209 CGTCCGCTCGCGGATGAAATGATCCATACGCTACTCTTGGATCCATGATGATTCGCCGA 268

QY 257 CGATGGATATGATGATTTATTCGACGAGAAATTAACAAATCATATAAGTATATGTTGAGT 316
DB 269 GCATGACACAGATGATTTGCGCGCGCAAGCGGACGACCAATTAAGTTCGCGGAGG 328

QY 317 GGACTCGGATATAGCAGGTGATGCTTTATTAATACTAAAGCATTTGAACCTATTTC----- 371
DB 329 CGATGGCCATCTTTGGCGGGGACGGTGTGACGTACGCGTTTCAATTGATCACCAGAA 388

QY 372 -AAGTGATGATAGATTAACCTGATGAAGTAAATAAATTAAGTTCTACACGCTCTCAATAG 430
DB 389 TCACATGATGAGCGATCCCTCTTCGTCGGCTTCGCGCTCATCGAACGCTGGCGAAG 448

QY 431 CRAAGTGTGATGTTGGAATGTTGCGCGGTCAATGTTAGATATGCAAGCGAAGGCAAC 490
DB 449 CGSCCGGTCCGGAAGGATGTTGCGCGGTGAGCGCGGATGATGAGGAGGAGGAA 508

QY 491 CAATGATCTGGAACCTTTGGAATGATACACAAAACAAAACAGGAGGATTAATTA 550
DB 509 CGTGTAGCTTTGCGAGCTCGATATATTCATCGGCATTAACCGGGGAAATGCTGCAAT 568

QY 551 TTGCGGTTATGAGTCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
DB 569 ACAGCTGTACGCGCGGCTTGATCGCGCGCTGATGCGCGGCAACCGCGGAGCTTG 628

QY 611 AAGTTATGTTATATATAGTATGATGTTCCAGATTTAAAGATGATTTATTAGACTGCT 670
DB 629 ACGAATTCGCGGCCATAGGCTTGCCTTTCAAAATTCGCGATGATATCTCGATTTG 688

QY 671 ATGCTGATGAACAAAGTTAGTTAAAGTGGCAGCGGCTTTGAAATATAAAGTA 730
DB 689 AAGGGCAGAAAGAAAATTCGGCAAGCGGTCGGCAGCGACCAACAAAGCGA 748

QY 731 CGTACCTGATTTATTAGGAAAGATGGCGCAGAGATAAATGACTTATCATAGAGAG 790
DB 749 CGTATCCAGGTTGCTGCTGCTGCGCGCGGAGGAAAGTGGGTTCCATATCGAGG 808

QY 791 CAGC 794
DB 809 CGGC 812

RESULT 14

US-09-217-609A-28
Sequence 28, Application US/09217609A
Patent No. 6071733
GENERAL INFORMATION:
APPLICANT: MURAMATSU, Masayoshi
APPLICANT: KOIKE, Ayumi
APPLICANT: OGURA, Kyoza
APPLICANT: KOYAMA, Tanetoshi
APPLICANT: SHIMIZU, Naoto
APPLICANT: CHO, Yenwin
TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, NW - Suite 600
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: WordPerfect 6.1 for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/217,609A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/873,235
APPLICATION NUMBER: 08/873,235
FILING DATE: 11-Jun-1997
ATTORNEY/AGENT INFORMATION:
NAME: TOFFENETTI, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 10235/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-09-217-609A-28

Query Match 17.5%; Score 150.6; DB 3; Length 486;

Best Local Similarity 58.4%; Pred. No. 5.2e-28;

Matches 277; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 121 GGAGGTAAAGCATCCGACCGCTTCTGTTATTACTCATTCTAGCTTACCTAATACCGAG 180
DB 1 GGTGGCAGCGCATTAAGCACTTACTTGTCTGACTACTTTAGATAGTTAGGTGGCAAT 60

QY 181 TATGAGTTAGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCACCTATT 240
DB 61 GCACATGACGGTTTACCATTGGCAATGGCTTGAATGATTCATACGTTATTCTTAATT 120

QY 241 CATGATGACCTACCGCGGATGATATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 121 CAGCATGACTTGGCGCAATGGAATGATGATGATGATGATGATGATGATGATGATGAT 180

QY 301 AAGATATATGTCAGTGGACTGCGATATAGCAGGTGATGCTTTATTAACTAAGCATTT 360
DB 181 AAGCGTTTGTAGAACCAACAGCTACTGCTGGAGATGCAATGCTCAGCTGATGCTTTT 240

QY 361 GAACCTTATTCAGTATGATGATTAAGTGAAGTAAATAAATAAAGTTCTTACACCG 420
DB 241 CAATGCAATTTAA--ATACGCGATTAAACGCAGAAATTAATATCATTTGATTAATTA 297

QY 421 CTGTCAATAGCAAGTGGTCAATGTTGGATGTTGGCGGCGTCAAAATGTTAGATGCAAGC 480
DB 298 TTAAGTACTGCTTCTGGATCTAATGCGCATGGTTTACGGCAAAATGCTCGATATGCAAGT 357

QY 481 GAAGGCCAACCAATTCATCTTGAACCTTTGGAATGATACACAAACAAACAGGCA 540
DB 358 GAACATAAACCAATTCATCTTGAACCTTTGGAATGATACACAAACAAACAGGCGAN 417

QY 541 TTATTAACTTTTTCGGGTTATGATGTCAGCAGATATCGCTAAATGTCGATGATACA 594
DB 418 TTGATTCGTCGACGANTTGTAAAGTCAGGATCATCATANTGANTTTTANTGATGCA 471

RESULT 15

US-08-873-235B-28
 ; Sequence 28, Application US/08873235B
 ; Patent No. 6174715
 ; GENERAL INFORMATION:
 ; APPLICANT: MURAMATSU, Masayoshi
 ; APPLICANT: KOIKE, Ayumi
 ; APPLICANT: OGURA, Kyoza
 ; APPLICANT: KOYAMA, Tanetoshi
 ; APPLICANT: SHIMIZU, Naoto
 ; APPLICANT: CHO, Yenwin
 ; TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kenyon & Kenyon
 ; STREET: 1025 Connecticut Avenue, NW - Suite 600
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: US
 ; ZIP: 20036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
 ; SOFTWARE: WordPerfect 6.1 for Windows
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/873,235B
 ; FILING DATE: 11-Jun-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 154441/1996
 ; FILING DATE: 14-Jun-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: TOFFENETTI, Judith L.
 ; REGISTRATION NUMBER: 39,048
 ; REFERENCE/DOCKET NUMBER: 10235/1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-429-1776
 ; TELEFAX: 202-429-0796
 ; INFORMATION FOR SEQ ID NO: 28:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 486 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: genomic DNA
 ; US-08-873-235B-28

Query Match 17.5%; Score 150.6; DB 4; Length 486;
 Best Local Similarity 58.4%; Pred. No. 5.2e-28;
 Matches 277; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

Qy	121	GGAGTAAACGACCGACGAGTCTCTGTATTACTCACTTTAGATTCACTAAATACCGAG	180
Db	1	GGTGGCAGCGCATAGACCATTACTTCTCTGACTACTTTAGATAGTTAGTGGCAAT	60
Qy	181	TATGAGTTAGTATGACAGCGCAATTGCACATAGAAATGATCATACATATTCATTATT	240
Db	61	GCACATGACGGTTTACCATTGGCATTGCGCTTGAATGATTTCATCGTATCTTTAAT	120
Qy	241	CATGATGACCTACCAGCGATGATTAATGATGATTCGACGAGAAATTAACAATCAT	300
Db	121	CACGATGACTTGGCGGAATGGATTAATGATGATTCGTCGCGTAAACTCACCAGTAT	180
Qy	301	AAAGTATATGTGAGTGCATGCGCATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT	360
Db	181	AAGCGTTTGTATGAGCAACAGCTATACTCGCTGGAGATGCTTGCCTGCTGATGCTTTT	240
Qy	361	GAACCTATTTCAGTATGATGATTAACATGATGATGATGATGATGATGATGATGATG	420
Db	241	CAATGCATTTTAA---ATACGCACTTAAACCGACAAATTAATTTATCATTCATTAA	297
Qy	421	CTGTCAATAGCAAGTGGTCATGTTGGAAATGGTCGCGCGGTCAATGTTAGATATGCA	480

Db	298	TTAAGTACTGCTTCTGGATCTAATGGCATGTTTACGGCCAAATCTCGATATGCAAGGT	357
Qy	481	GAAGGCCAACCAATTGATCTTGAACACTTTGGAATGATACACAAAAACAAACAGGAGCA	540
Db	358	GAACATAAACACATTGACATTAATGAACCTGGACGTATTACATACATAAAACCGGTGAN	417
Qy	541	TTATTAACCTTTGGCGTTTATGAGTGCAGCAGATATCGCTAATGTGCTGATGATACA	594
Db	418	TTGATTCTGTCAGCAGTAAAGTGCAGGTATCATATGANTTTTANTGATGCA	471

Search completed: May 30, 2003, 17:36:21
 Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 15:24:59 ; Search time 1600 Seconds
(without alignments)
8715.196 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatcaccgatgaa.....tattagaatcgttgattta 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
tal number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151.6	17.6	4484	17 BH770981	BH770981 LLMGtag70
2	89.2	10.4	354	12 BF610197	BF610197 NXSI_055_
3	89.2	10.4	558	9 AL749888	AL749888 AL749888
4	88.6	10.3	657	9 AJ502290	AJ502290 AJ502290
5	88	10.2	425	10 BE323612	BE323612 NXF006A09P
6	87.4	10.2	592	12 BG317831	BG317831 NXPV_006_

7	84.6	9.8	355	9	AI780894
8	84.2	9.8	442	9	AJ302129
9	80	9.3	571	10	BE432698
10	79.2	9.2	475	17	BH439236
11	78.6	9.1	468	12	BF519441
12	78.6	9.1	582	14	BQ56404
13	78.4	9.1	685	13	BQ56404
14	77	8.9	636	9	AU238772
15	75.4	8.8	716	14	BQ506776
16	75.2	8.7	863	17	BH584613
17	75	8.7	770	17	BH250143
18	74.8	8.7	808	14	BQ506777
19	73.4	8.5	384	10	BE521431
20	73	8.5	703	17	BH429531
21	70.6	8.2	1101	17	CNS0039G
22	70.4	8.2	561	12	BF054034
23	70.4	8.2	630	12	BF270259
24	70.4	8.2	670	14	BQ401843
25	70.4	8.2	689	12	BG439911
26	70.4	8.2	892	12	BF276100
27	70.4	8.2	907	12	BG444779
28	69.8	8.1	701	10	BE238154
29	69.6	8.1	666	13	BI717621
30	69.6	8.1	575	13	BI723806
31	68.8	8.0	668	13	BI924206
32	68.8	8.0	669	12	BG889977
33	68.8	8.0	681	10	AW096600
34	68.8	8.0	712	12	BF521402
35	68	7.9	668	13	BI726915
36	67.6	7.9	838	17	BH657017
37	67.2	7.8	709	17	BH502634
38	67	7.8	669	17	BH731669
39	67	7.8	823	10	BE640755
40	66.2	7.7	542	14	BQ696145
41	64.2	7.5	662	10	AW725795
42	64	7.4	828	17	BH501300
43	64	7.4	852	17	BH432244
44	62	7.2	618	9	AU238793
45	62	7.2	706	13	BI434201

ALIGNMENTS

RESULT 1
BH770981
LOCUS BH770981 484 bp DNA linear GSS 01-MAY-2002
DEFINITION LLMGtag706 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, DNA sequence.
ACCESSION BH770981
VERSION BH770981.1 GI:20373938
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Lactococcus lactis subsp. cremoris.
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 4484)
AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL Sci. Aliments, (2002) In press
COMMENT Contact: Sorokin A
Genetique Microbienne
INRA
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is xsea (94%)
Class: shotgun
High quality sequence start: 30
High quality sequence stop: 4456.
Location/Qualifiers
1. .4484

FEATURES
source

/organism="Lactococcus lactis subsp. cremoris"
/strain="MG1363"
/db_xref="taxon:1359"
/clone_lib="MG1363 Random Sequence Tag Library"
/note="Vector: pSGM2; Site_1: SmaI; Library of
chromosomal fragments of L.lactis strain MG1363 was
prepared by partial AluI digestion or by sonication."
BASE COUNT 1540 a 637 c 1023 g 1284 t
ORIGIN

Query Match 17.6%; Score 151.6; DB 17; Length 4484;
Best Local Similarity 52.8%; Pred. No. 1.9e-26;
Matches 378; Conservative 0; Mismatches 334; Indels 6; Gaps 2;
QY 106 TATTCATTAAATGCTGGAGTAACGATCCGACCAAGTTCTGTATTACTCACTTAGAT 165
Db 3738 TATTCCTTAATGCGGCGGTAAACGAATTCGCTCTCTCTTTTGTGATCTTTGGAA 3797
QY 166 TCACTAAATACCGAGTAT---GAGTTAGGTATGAAGAGCGCAATTCGACTAGAAATGAT 222
3798 GCTTTTGATTTGAAGTTACAACGGCCCATTTATCAAGTGGCCGCGCTTGAATGAT 3857
QY 223 CATACATATTCATTTATCATGATGACCTACGAGCGATGATATGATGATATTCGACGA 282
Db 3858 CATACGGGCTCGTAAATTCATGATGACTTGCCTGCGATGGACAAATGATGATATCGTGT 3917
QY 283 GGAATAATTAAACAAATCATAAAGTATATGTTGAGTGCAGCTGCGATATTAGCAGGTGATGCT 342
Db 3918 GGAAGTTACGAACCAATAAANATTTGATGAAGCGACAGCAATTCGCTGGAGTACT 3977
QY 343 TTATTAACTAAAGCACTTAATTTCAAGTGATGATGATAGATTAAGTGAAGTAAAT 402
Db 3978 TTATTTTTCGACCTTTTATGTTCTTCAAGTAGCGAT---TTGTCGGCGCAACAAT 4034
QY 403 ATAAAGTTCTACACGGCTGTCATAGCAAGTGGTCATGTTGGATGTTGCGCGGTCAA 462
Db 4035 GTATCATTTGACCGCGGAATTTGGCTTATGCTCGGCTCGTATGGAATGTTGCTGTCAA 4094
QY 463 ATGTTAGATATGAAAGCGAAGCCCAACCAATTTGATCTTTGMAACTTTGGAAATGATACAC 522
Db 4095 ATTTGGAATGCTGTTGAGGTAAAGATTTAACTTTGACTCAGATTGACCAATTCAT 4154
QY 523 AAACAAAACAGGAGCATTTTAATTTTGGGGTTATGAGTGACGACAGATATCGGTAAT 582
Db 4155 CGATTAAACCGGTCGACTTTTAATTTTCTTTTGTGCTGCTGGAATCGTAGCCCAA 4214
QY 583 GTCGATGATACAACTAAAGCAATTTAGAAAGTTATAGTTATCATTTAGTATGTTTC 642
Db 4215 AAACAAATCAAGAAATCGAAAACTTCGTGTAGTCGGACAAATTTTAGATGATGCTTTT 4274
QY 643 CAGATTAAAGTATGATTTATTAGACTGCTATGTTGATGAAGCAAAAGTTAGGTAAAAAGTG 702
Db 4275 CAAATTCGTGATGACATTTTAGATGTGACGGCCACTTTTTCGTAACCTGGTAAGCGCCT 4334
QY 703 GCGAGCGCTGTAATAATTAAGTACGTACGTGAGTTTATTAGGGAAGATGCGCGCA 762
Db 4335 GGAAGGATGTTTGAAGAAAGTGCAGTTATGTTGCGCAATTTAGGACTTGAAGAGCT 4394
QY 763 GAAGATAAATGACTTTATCATAGAGACGCGAGTGGATGAATCAACGCAAAATTTGATG 820
Db 4395 AAATATCACTGACAGATAAGTTGTCAGAAGTAAAAAATTTACTGACAGATTGAATG 4452

RESULT 2
BF610197
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BF610197
NXSI_055_H03_F NXSI (Nsf xylem side wood inclined) Pinus taeda cDNA
clone NXSI_055_H03_5', mRNA sequence.
BF610197
BF610197.1 GI:11778609
EST.
loblolly pine.
Pinus taeda

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.
Sederoff, R.
Molecular Basis of Wood Formation in the Pine Megagenome
Unpublished (2000)
Contact: Johnson, Arthur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801
Email: ajohnson@unity.ncsu.edu
Seq primer: T3.
Location/Qualifiers
1. .354
/organism="Pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXSI_055_H03"
/clone_lib="NXSI (Nsf xylem side wood inclined)"
/tissue_type="xylem"
/cell_type="juvenile"
/dev_stage="juvenile"
/lab_host="XLI-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form side
wood by bending to a 45 degree angle and tying them to the
ground. Differentiating xylem was harvested from the sides
of the inclined stems, and a mixture of all three
genotypes was used for the library. oligo-dt primed cDNA
was directionally cloned into the EcoRI-XhoI Bluescript SK
vector arms. NOTE: The sequences contain a 'cDNA adapter'
between the EcoRI site and the start of the Est. The
adapter sequence is 'AATTCGACGAG'." 104 t 12 others
BASE COUNT 72 a 68 c 98 g 104 t 12 others
ORIGIN

Query Match 10.4%; Score 89.2; DB 12; Length 354;
Best Local Similarity 63.6%; Pred. No. 1.9e-11;
Matches 136; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 182 ATGAGTTAGTATGAAGCGCAATTTGCACATAGAAATGATTACATATTCCTTATTC 241
Db 49 AGGAGATGGCAATGCCAGCTGCTGTGCAATGAGATTATCCACACAAATGCTTTGATTC 108
QY 242 ATGATCACTTACGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
Db 109 ATGATGATGCTGCTTCCATGGATGATGATGATGATGATGATGATGATGATGATGATGAT 168
QY 302 AAGTATATGCTGAGTGGCTGCGATATTAGCAGGTGATGCTTTTATTAAGTAAAGCATTTG 361
Db 169 AGGTCTTTGTTGAGGCGCACTGCGAGTTCTTGCAGTGTGCTCTCTCTCTCTCTCTCT 228
QY 362 AACTTATTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 395
Db 229 AGCACATTGCGTGGCTACAAGCAAGCAAGTGTGGA 262

RESULT 3
AL749888
LOCUS

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
AL749888
AL749888 AS Pinus pinaster cDNA clone AS02C09 similar to
GERANYLGERANYL DIPHOSPHATE SYNTHASE, mRNA sequence.
AL749888.1 GI:21491122
EST.
Pinus pinaster.
Pinus pinaster
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.
Frigerio, J. and Plomion, C.

Contact: Harrison MJ
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7325
Fax: 580 221 7380
Email: mjharrison@noble.org
Medicago Genome Initiative accession: MGI:S:20177
Insert Length: 818 Std Error: 0.00
Plate: 006 row: A column: 09
Seq primer: TCACACGAGGAACAGCTATGAC.

FEATURES

source
1. 425
Location/Qualifiers
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF006A09PL"
/clone_lib="Phosphate starved leaf"
/tissue_type="leaf"
/note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20uM potassium phosphate. RNA was prepared from above ground tissues."
BASE COUNT 109 a 116 c 96 g 104 t

Query Match 10.2%; Score 88; DB 10; Length 425;
Best Local Similarity 57.1%; Pred. No. 4e-11;
Matches 160; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 92 AAGAAAGTATGTTGTTATTCATTAATGCTGGAGGTAAAGCATCCGACCAAGTCTCTGTAT 151
DB 124 ATGAGCCATGCTTACTCTCTCTCGCGGGGAAACGCGTTCGCGCGGTTCTCTGT 183
QY 152 TACTCACTTAGATTCACTAATACCGAGTATGAGTTAGTAGTGAAGAGCGCAATTCGAC 211
DB 184 TAGCCGCTGTGAACCTCGTGGAGAACCGACCGATGCGGATGCCAGCGCTCGCGCC 243
QY 212 TAGAAATGATTCATACATATTCATTAATGCTGGAGGTAAAGCATCCGACCAAGTCTCTGTAT 271
DB 244 TTGAATGATCCACAGTGTCTCTCATCCAGTATGACCTCTCTGATGGATGAAGATG 303
QY 272 ATTATCGACGAGAAATTAACAATATCAAAATATATGATGAGTGGACTGCGCATATTAG 331
DB 304 ATCTCCGGCGAGGTAAACCTACAACACCAAAAGTCTTCGAGAGAGAGCTTGCTGTTCTCG 363
332 CAGGTGATGCTTATTAACATAAGCAATTTGAACCTATTTC 371
364 CGGAGATGCTCTCTCTCGCTTTTGTCTTTCGAACATATTGC 403

RESULT 6
BG317831
LOCUS
DEFINITION
NXPV_006_B05_F NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda
cdna clone NXPV_006_B05 5', mRNA sequence.
ACCESSION
BG317831
VERSION
BG317831.1 GI:13127261
KEYWORDS
EST.
SOURCE
loblolly pine.
ORGANISM
Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE
1 (bases 1 to 592)
Sederoff, R.
Molecular Basis of Wood Formation in the Pine Megagenome
Unpublished (2000)
Contact: Johnson, Arthur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801

Email: ajohnson@unity.ncsu.edu
Seq primer: T3.
Location/Qualifiers
1. 592

FEATURES

source
/organism="Pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXPV_006_B05"
/clone_lib="NXPV (Nsf Xylem Planings wood Vertical)"
/tissue_type="Xylem"
/cell_type="Planings (secondary)"
/lab_host="XLI-Blue"
/note="Vector: Bluescript SK; Site1: Eco RI; Site2: XhoI
from a ten year old tree in the transitional phase. The tree is a kind gift of the Westvaco Corporation. Secondary xylem was harvested from the tree by peeling back the bark and primary xylem and then removing the underlying tissue with a block plane. NOTE: The sequences contain a 'cdna adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTGGGACGAG'."
BASE COUNT 163 a 118 c 134 g 154 t 23 others

Query Match 10.2%; Score 87.4; DB 12; Length 592;
Best Local Similarity 53.9%; Pred. No. 5.9e-11;
Matches 151; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 92 AAGAAAGTATGTTGTTATTCATTAATGCTGGAGGTAAAGCATCCGACCAAGTCTCTGTAT 151
DB 304 ATGAGCAATAGGATTCCTCTGCTGGGGGAAAGCGTGTGAGACCATATTATGCA 363
QY 152 TACTCACTTAGATTCACTAATACCGAGTATGAGTTAGTAGTGAAGAGCGCAATTCGAC 211
DB 364 TCCTCTGCGATCTCGTAGGTGGTCTGAGGAATCAGTCATGCCATCGGCTGTGCAA 423
QY 212 TAGAAATGATTCATACATATTCATTAATGCTGGAGGTAAAGCATCCGACCAAGTCTCTGTAT 271
DB 424 TTGAGATGATTCACAAATGCTCTGATTCATCATGATTTGCCCTGTATGGCAATGACG 483
QY 272 ATTATCGACGAGAAATTAACAATATCAAAATATATGATGAGTGGACTGCGCATATTAG 331
DB 484 ATTTGAGAGNAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 543
QY 332 CAGGTGATGCTTATTAACATAAGCAATTTGAACCTATTTC 371
DB 544 CNGNNATGCTTTGTTNNCTTTTGCCTTNNAGCAATTC 583

RESULT 7
AI780894
LOCUS
DEFINITION
EST261773 tomato susceptible, Cornell Lycopersicon esculentum cdna
clone CUES13B2, mRNA sequence.
ACCESSION
AI780894
VERSION
AI780894.1 GI:5278935
KEYWORDS
EST.
SOURCE
tomato.
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 355)
D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Opton, J., Renning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D., Giovannoni, J.J., and Martin, G.B.
Generation of ESTs from Pseudomonas susceptible tomato
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute

Best Local Similarity 50.5%; Pred. No. 5.7e-09;
Matches 192; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY	106	TATTCATTAAATGCTGGAGGTTAAACCGATCCGACCGAGTCTGTGTTATCTACTCACTTTAGAT	165
Db	474	TACTCTCTCTGCGCGCGGGAAGCGAGTGAGACCGGTCTCTGCATCGCGCGGTGCGAG	415

QY
Dh

166 TCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACATAGAAATGATTTCAT 225
414 CTGGTGCCGCCGACACAGACTCCTCCCTCTCTTTTCCGCCCCGGCTTCCGGCCCCGCTTCTGTCAGC 255

[illegible]

Accession	Gene	Accession	Gene
Db	354	QY	286

Db	294	ANACCACCAACCAAGGTTCGGCGAGACGTCGGGTTTTAGCCGGAGACCGGTT	235
Qy	346	TTAACTAAAGCATTTGAACCTATTTCAGGTGATGATAGATTAACTGATGAAGTAAATA	405

Db	234	TTATCGTT	CGGTTT	IIIIII	IIIIII	CGGTCG	AGCTCG	GGGGT	GGCTCC	GGCGAG	GGTGGT	175
Qv	406	AAAGTTC	TAAC	ACGGT	CTC	TAAT	TGAC	TGCT	CAAT	TCG	CAAT	455

Db

174 AGAGCGATCGGGAGCTGGCGAGAGCGCTTGGATCGAAGAGCGCTTGTGGCGGGTCAGATT 115

Qy	466	TTAGATATGCAAAGCGAAGG	485
Db	114	GTGGATATCAGCAGCGAAGG	95

RESULT 11
BF519441

LOCUS	BF519441	468 bp	linear	EST 08-DEC-2000
DEFINITION	DSIL Medicago truncatula cDNA clone pDSIL-20N21, mRNA			
ACCESSION	BF519441			

VERSION B519441.1
 EST. barrel medic
 KEYWORDS
 SOURCE GI:11608124

ORGANISM

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosales; eucrosid 1; Fabales; Fabaceae; Papilionoideae; subfamily

REFERENCE

1 (bases 1 to 468)

AUTHORS

Fedorova, M., Pierson, B.L., Samac, D.A., Vance, C.P., Gantt, G.S., peng

Medicag. cynosucl. s., fabales, fabaceae, papilionoideae; trifoliolate

TITLE	Author
ESTS from leaves of <i>Medicago truncatula</i> after inoculation with <i>Callotetranychus trifolii</i>	Holt, I. E. and Fraser, C. M.
ESTs from leaves of <i>Medicago truncatula</i> after inoculation with <i>Callotetranychus trifolii</i>	H., Ellis, L., Town, C. D., Bowman, C. L., Craven, M. B., Hansen, T. S.,

JOURNAL COMMENT
 Published online 11 October 2001
 Unpublished (2000)
 Contact: Deborah A. Samac
 Department of Plant Pathology

University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058

Email: debbys@puccini.crl.umn.edu
University of Minnesota name: M276873e
TIGR sequence name: MTF683TK

FEATURES
source 1 458
Location/Qualifiers
Seq primer: SKmod (CTA GAA CTA gtg gAT CC).
More information is available at: <http://chrysie.tamu.edu/medicago>

```

/organism="Medicago truncatula"
/cultivar="genotype Al7"
/db_xref="taxon:3880"

```

```

/clone="pDSil-20N21"
/clone.lib="DSIL"
/tissue_type="leaves infected with Colletotrichum
+trifolii"

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:47:05 ; Search time 35 Seconds
(without alignments)
1092.655 Million cell updates/sec

Title: US-09-925-637-64

Perfect score: 1453

Sequence: 1 MTNLPKMLIDEVNNELSA.....ELTQIDEQFNKHLLEIVDL 287

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1453	100.0	287	22 AAU00859	S. aureus Farnesyl
2	1442	99.2	293	22 AAU36990	Staphylococcus aur
3	1436	98.8	293	21 AAB23333	Staphylococcus aur
4	1415	97.4	288	22 AAU33743	Staphylococcus aur
5	1015	69.9	302	23 ABP38326	Staphylococcus epi
6	943	64.9	260	22 AAG81565	S. epidermidis ope
7	702.5	48.3	293	23 ABB48306	Listeria monocytog
8	657	45.2	297	17 AAU00285	Mutant farnesylidp
9	656	45.1	297	17 AAU00286	Native farnesylidp
10	654	45.0	297	19 AAU62532	Farnesyl diphosphat

11	654	45.0	297	21 AAB29399	Bacillus stearothe
12	653	44.9	297	17 AAU00283	Mutant farnesylidp
13	650	44.7	297	17 AAU00284	Mutant farnesylidp
14	650	44.7	297	19 AAU47444	Bacillus stearothe
15	650	44.7	297	20 AAU27008	Farnesyl diphospha
16	647	44.5	297	14 AAR35047	FPS, Bacillus ste
17	646	44.5	297	20 AAU27006	Geranyl diphosphat
18	645	44.4	297	17 AAU00282	Mutant farnesylidp
19	644	44.3	297	20 AAU27007	Geranyl diphosphat
20	637	43.8	297	19 AAU62535	Mutant farnesyl di
21	637	43.8	297	19 AAU62537	Mutant farnesyl di
22	636	43.8	297	19 AAU62533	Mutant farnesyl di
23	618	42.5	295	19 AAU62536	Mutant farnesyl di
24	618	42.5	295	19 AAU62546	Mutant farnesyl di
25	617	42.5	295	19 AAU62534	Mutant farnesyl di
26	596	41.0	293	22 AAU35037	Enterococcus faeca
27	551.5	38.0	290	23 ABP25554	Streptococcus poly
28	551	37.9	291	21 AAB26267	Streptococcus pneu
29	549	37.8	291	22 AAU38004	Farnesyl diphospha
30	547	37.6	291	22 AAU37680	Streptococcus pneu
31	540.5	37.2	285	23 ABB54173	Lactococcus lactis
32	521	35.9	393	21 AAU82667	Taxus GPPP synthas
33	520.5	35.8	393	21 AAU82651	Taxus canadensis g
34	520.5	35.8	393	21 AAU82653	Taxus GPPP synthas
35	520.5	35.8	393	21 AAU82654	Taxus GPPP synthas
36	520.5	35.8	393	21 AAU82655	Taxus GPPP synthas
37	520.5	35.8	393	21 AAU82656	Taxus GPPP synthas
38	520.5	35.8	393	21 AAU82657	Taxus GPPP synthas
39	520.5	35.8	393	21 AAU82658	Taxus GPPP synthas
40	520.5	35.8	393	21 AAU82659	Taxus GPPP synthas
41	520.5	35.8	393	21 AAU82660	Taxus GPPP synthas
42	520.5	35.8	393	21 AAU82661	Taxus GPPP synthas
43	520.5	35.8	393	21 AAU82662	Taxus GPPP synthas
44	520.5	35.8	393	21 AAU82663	Taxus GPPP synthas
45	520.5	35.8	393	21 AAU82664	Taxus GPPP synthas

ALIGNMENTS

RESULT 1

AAU00859

ID AAU00859 standard; Protein; 287 AA.

XX AAU00859;

XX 04-JUL-2001 (first entry)

XX S. aureus Farnesyl diphosphatesynthase, Ispa.

DE Farnesyl diphosphatesynthase; Ispa; immunogen; vaccine; antibody;

KW wound infection; cellulitis; burn infection; eyelid infection;

KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;

KW skin infection; scalded skin syndrome; toxic epidermal necrosis;

XX Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis.

OS Staphylococcus aureus.

XX

FT Key Location/Qualifiers

FT Region 88..91

FT /label= Antigenic_epitope

FT Region 93..95

FT /label= Antigenic_epitope

FT Region 240..243

FT /label= Antigenic_epitope

FT WO200116292-A2.

XX

PN 08-MAR-2001.

XX

PD 31-AUG-2000; 2000WO-US23773.

XX

XX 01-SEP-1999; 99US-0151933.

XX

XX (HUMA-) HUMAN GENOME SCI INC.
XX Choi GH;
XX WPI: 2001-183259/18.
XX N-PSDB; AAS00821.
XX
XX New isolated nucleic acid for use in diagnosing Staphylococcus
XX infections and in vaccines for eliciting immune responses to the
XX infections -
XX
XX Claim 9; Page 23; 225pp; English.
XX
XX The sequence represents S. aureus IspA (Farnesyl diphosphatesynthase).
XX The polynucleotides of the invention are used to detect Staphylococcus
XX nucleic acids in a biological sample from an animal for diagnosing
XX Staphylococcus infections. The polypeptides of the invention are used to
XX detect anti-Staphylococcus antibodies in a biological sample from an
XX animal to diagnose Staphylococcus infections. The polypeptides are also
XX used in vaccines to elicit protective antibodies in an animal to a member
XX of the Staphylococcus genus and for preventing or attenuating an
XX infection caused by a member of the Staphylococcus genus e.g wound
XX infection, cellulitis, burn infection, eyelid infection, food poisoning,
XX joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
XX scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
XX disease and Lyell's disease), toxic shock syndrome and endocarditis. The
XX polynucleotides may also be used in vaccines and for preventing or
XX attenuating a Staphylococcus infection. Antibodies to the polypeptides
XX may be used to purify, detect and target the polypeptides in vitro and
XX in vivo diagnostic and therapeutic methods.

XX Sequence 287 AA;
XX
XX Query Match 100.0%; Score 1453; DB 22; Length 287;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-125; Indels 0; Gaps 0;
XX Matches 287; Conservative 0; Mismatches 0;
XX
XX QY 1 MTNLPNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLILLTDSLNTE 60
XX Db 1 MTNLPNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLILLTDSLNTE 60
XX
XX QY 61 YEIGMKSALAEIMHTYSLIHDDLPAMDNDYRGLTNHVKVGEWTAIAGDALLTKAF 120
XX Db 61 YEIGMKSALAEIMHTYSLIHDDLPAMDNDYRGLTNHVKVGEWTAIAGDALLTKAF 120
XX
XX QY 121 ELISSDRLTDEVKIKVQLRSLASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTGA 180
XX Db 121 ELISSDRLTDEVKIKVQLRSLASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTGA 180
XX
XX QY 181 LLTFVAVMSAADIANVDTTKEHLESYSYHLGMMFQIKDDLLDCYGDGAKLGGKVGSDLEN 240
XX Db 181 LLTFVAVMSAADIANVDTTKEHLESYSYHLGMMFQIKDDLLDCYGDGAKLGGKVGSDLEN 240
XX
XX QY 241 NKSTYVSLGKGAEDKLTTHRDAAVDELQIDQFNTKHLLEIVDL 287
XX Db 241 NKSTYVSLGKGAEDKLTTHRDAAVDELQIDQFNTKHLLEIVDL 287

RESULT 2

AAU36990

ID AAU36990 standard; Protein; 293 AA.

XX

XX AAU36990;

XX 14-FEB-2002 (first entry)

XX

XX Staphylococcus aureus cellular proliferation protein #1160.

XX

XX Antisense; prokaryotic cellular proliferation protein;

XX

XX antibiotic; antibacterial; drug design.

XX

XX Staphylococcus aureus.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

WO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US09180.

21-MAR-2000; 2000US-191078P.

23-MAY-2000; 2000US-206848P.

26-MAY-2000; 2000US-207272P.

23-OCT-2000; 2000US-242578P.

27-NOV-2000; 2000US-253625P.

22-DEC-2000; 2000US-257931P.

16-FEB-2001; 2001US-269308P.

(ELIT-) ELITRA PHARM INC.

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

Yamamoto RT, Xu HH;

WPI: 2001-611495/70.

N-PSDB; AAS34849.

New polynucleotides for the identification and development of

antibiotics, comprise sequences of antisense nucleic acids -

Example 3; Seq ID No 12583; 511pp; English.

The invention relates to antisense inhibitors of genes essential to

prokaryotic cellular proliferation, their use in identifying the

genes, their use in the discovery of novel antibiotics, the essential

genes themselves and the encoded proteins. The prokaryotes used are

Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

invention is also useful for the identification of potential new targets

for antibiotic development. The antisense nucleic acids can also be used

to identify proteins used in proliferation, to express these proteins,

and to obtain antibodies capable of binding to the expressed proteins.

The proteins can be used to screen compounds in rational drug discovery

programmes. The antisense nucleic acid sequence is also useful to screen

for homologous nucleic acids which are required for cell proliferation in

a wide variety of organisms. The present sequence represents an

essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part

of the printed specification, but was obtained in electronic

format directly from WIPO at

ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 293 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 99.2%; Score 1442; DB 22; Length 293;

Best Local Similarity 99.3%; Pred. No. 4.4e-124;

Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTNLPNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLILLTDSLNTE 60

Db 1 MTNLPNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLILLTDSLNTE 60

QY 61 YEIGMKSALAEIMHTYSLIHDDLPAMDNDYRGLTNHVKVGEWTAIAGDALLTKAF 120

Db 61 YEIGMKSALAEIMHTYSLIHDDLPAMDNDYRGLTNHVKVGEWTAIAGDALLTKAF 120

QY 121 ELISSDRLTDEVKIKVQLRSLASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTGA 180

Db 121 ELISSDRLTDEVKIKVQLRSLASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTGA 180

QY 181 LLTFVAVMSAADIANVDTTKEHLESYSYHLGMMFQIKDDLLDCYGDGAKLGGKVGSDLEN 240

Db 181 LLTFVAVMSAADIANVDTTKEHLESYSYHLGMMFQIKDDLLDCYGDGAKLGGKVGSDLEN 240

QY 241 NKSTYVSLGKGAEDKLTTHRDAAVDELQIDQFNTKHLLEIVDL 287

Db 241 NKSTYVSLGKGAEDKLTTHRDAAVDELQIDQFNTKHLLEIVDL 287

RESULT 3

AAB23333 standard; Protein; 293 AA.

AC AAB23333;

DT 12-JAN-2001 (first entry)

XX Staphylococcus aureus IsPa.

XX IsPa: bacterial disease; respiratory tract infection;

KW gastrointestinal infection; cardiac infection; Helicobacter pylori;

KW stomach cancer; stomach ulcer; gastritis.

XX Staphylococcus aureus.

OS US6107058-A.

PN 22-AUG-2000.

XX 26-MAR-1999; 99US-0276873.

PR 26-MAR-1999; 99US-0276873.

XX (SMIK) SMITHKLINE BEECHAM CORP.

PA Gwynn M, Wilding EI;

PI WPI; 2000-578535/54.

DR N-PSDB; AAA92031.

XX Novel farnesyl diphosphate synthase polynucleotide from staphylococcus aureus useful for diagnosis and treatment of bacterial infections and as hybridization probe for isolating genomic clones -

PS Claim 13; column 3-4; 15pp; English.

The present sequence is the Staphylococcus aureus IsPa protein. This protein and its coding sequence can be used in many research assays, as well as treatments for bacterial diseases such as infections of the respiratory tract (including otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis, empyema and lung abscesses), cardiac infections such as infective endocarditis, gastrointestinal infections including secretory diarrhoea, splenic abscesses and retroperitoneal abscesses, CNS infections such as cerebral abscesses, eye infections (including blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis and dacryocystitis), kidney and urinary tract infections such as epididymitis, intrarenal and perinephric abscesses and toxic shock syndrome, skin diseases (including impetigo, folliculitis, cutaneous abscesses, wound infection and bacterial myositis), bone and joint infections such as septic arthritis and osteomyelitis, septic thrombophlebitis, food poisoning and scalded skin syndrome. In addition, they can be used to treat diseases caused by Helicobacter pylori, including stomach cancer, stomach ulcers and gastritis.

XX Sequence 293 AA;

Query Match

Best Local Similarity 98.8%; Score 1436; DB 21; Length 293;

Matches 284; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTNLPKMLIDEVNNELSVAINKSVMDTQLESMYSLNAGKRIRPVLLLTLDLSNTE 60

DB 1 MTNLPKMLIDEVNNELSVAINKSVMDTQLESMYSLNAGKRIRPVLLLTLDLSNTE 60

QY 61 YELGKSAIALEMIHTYSLIHDDLPAMDNDYRRKLTNHNKYGWTAIAGDALLTTF 120

DB 61 YELGKSAIALEMIHTYSLIHDDLPAMDNDYRRKLTNHNKYGWTAIAGDALLTTF 120

QY 121 ELISSDDRLTDEVKIKVLRSLIASGHVGMVGQMLDMQSEGQPIDLETLEMIHKTGA 180

DB 121 ELISSDDRLTDEVKIKVLRSLIASGHVGMVGQMLDMQSEGQPIDLETLEMIHKTGA 180

QY 181 LITFAVMSAADIANVDDTTKEHLESYSYHLGMFMFQIKDDLDCYGDGKLGKVGSDLEN 240

DB 181 LITFAVMSAADIANVDDTTKEHLESYSYHLGMFMFQIKDDLDCYGDGKLGKVGSDLEN 240

QY 241 NKSTVSLLGKGDGAEKLTTHRDAAVDELDTQIDQFNTHKLLLEIVDL 287

DB 241 NKSTVSLLGKGDGAEKLTTHRDAAVDELDTQIDQFNTHKLLLEIVDL 287

RESULT 4

AAU33743

ID AAU33743 standard; Protein; 288 AA.

XX AAU33743;

XX 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #19.

XX Antisense; prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

OS WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

PA Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS51602.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 5239; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 288 AA;
Query Match 97.4%; Score 1415; DB 22; Length 288;
Best Local Similarity 99.3%; Pred. No. 1.3e-121;
Matches 280; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSLNTYELGM 65
DB 1 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSLNTYELGM 60
QY 66 KSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHHKVGWTAIAGDALLTKAFELISS 125
DB 61 KSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHHKVGWTAIAGDALLTKAFELISS 120
QY 126 DDLRTDEVKIKVQLRSLIASGHVGMVGQMDMQSEGQPIDLETLEMIHKTGTGALLTFA 185
DB 121 DDLRTDEVKIKVQLRSLIASGHVGMVGQMDMQSEGQPIDLETLEMIHKTGTGALLTFA 180
QY 186 VMSAADIANVDDTTKEHLESYSYHLGMFMFQIKDDLDCYGDGDEAKLGKVGSDLENKSTY 245
DB 181 VMSAADIANVDDTTKEHLESYSYHLGMFMFQIKDDLDCYGDGDEAKLGKVGSDLENKSTY 240
QY 246 VSLLGKDGAEKLTTHRDAAVDELTDQIDQFNTHKLLLEIVDL 287
DB 241 VSLLGKDGAEKLTTHRDAAVDELTDQIDQFNTHKLLLEIVDL 282
RESULT 5
ABP38326
ID ABP38326 standard; Protein; 302 AA.
XX AC ABP38326;
XX DT 24-JUL-2002 (first entry)
XX DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3171.
XX KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX KW antibacterial; gene therapy.
XX OS Staphylococcus epidermidis.
XX PN US6380370-B1.
XX PD 30-APR-2002.
XX PS 13-AUG-1998; 98US-0134001.
XX PR 14-AUG-1997; 97US-055779P.
XX PS 08-NOV-1997; 97US-064964P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm LA, Bush D;
XX PS WPI; 2002-381255/41.
XX DR N-PSDB; ABN90871.
XX PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
XX PT polypeptide, useful for diagnosing and treating bacterial infections -
XX PS Disclosure; SEQ ID 3171; 267pp; English.
XX CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX CC antibacterial activity and can be used in gene therapy. The sequences
XX CC can also be used in the diagnosis and treatment of bacterial infections,
XX CC particularly S. epidermidis infections. The sequences can be used to
XX CC screen for compounds able to interfere with the S. epidermidis life
XX CC cycle or inhibit S. epidermidis infection.
XX CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX SQ Sequence 302 AA;
Query Match 69.9%; Score 1015; DB 23; Length 302;
Best Local Similarity 68.6%; Pred. No. 6.8e-85;
Matches 197; Conservative 41; Mismatches 49; Indels 0; Gaps 0;
QY 1 MTNLPNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSLNT 60
DB 10 MKKLQMKLLINIINTSLNKSIQSPKTNLEESMKYSLNAGGKRIRPVLLLTLMKLNKD 69
QY 61 YELGMSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHHKVGWTAIAGDALLTKAF 120
DB 70 YQOGLSALALEMIHTYSLIHDDLPAMDNDYRRGKLTNHHKVGWTAIAGDALLTKAF 129
QY 121 ELISSDRLTDEVKIKVQLRSLIASGHVGMVGQMDMQSEGQPIDLETLEMIHKTGTGA 180
DB 130 ELYSNDTTIEDSVKVSIIKRLSKASGHLGMVGQALDMESEKSIKRLTESIHETKGA 189
QY 181 LITFAVMSAADIANVDDTTKEHLESYSYHLGMFMFQIKDDLDCYGDGDEAKLGKVGSDLEN 240
DB 190 LLNFSVMAAVDITQAEQNTAKNLDSESHLGMFMFQIKDDLDCYGDGDEAKLGKVGSDIVN 249
QY 241 NKSTYVSLGKDGAEKLTTHRDAAVDELTDQIDQFNTHKLLLEIVDL 287
DB 250 HKSTYVSLGKDGAEKLTTHRDAAVDELTDQIDQFNTHKLLLEIVDL 296
RESULT 6
AAG81565
ID AAG81565 standard; Protein; 260 AA.
XX AC AAG81565;
XX DT 03-SEP-2001 (first entry)
XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:224.
XX KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
XX KW vaccination; endocarditis.
XX OS Staphylococcus epidermidis.
XX PN WO200134809-A2.
XX PD 17-MAY-2001.
XX PS 09-NOV-2000; 2000WO-US30782.
XX PR 09-NOV-1999; 99US-0164258.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Kimmerly WJ;
XX PS WPI; 2001-316495/33.
XX DR N-PSDB; AAG52415.
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX PT useful for vaccinating against infections, e.g. endocarditis -
XX PS Claim 18; Page 103; 2188pp; English.
XX CC AAG52304 to AAG53970 represent nucleic acids (I) encoding polypeptides
XX CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
XX CC (I) and (II) can have antibacterial activity and therefore can be used
XX CC in vaccination. The nucleic acids (I) may be used to produce the
XX CC S. epidermidis polypeptides (II) via the production of vectors
XX CC containing them which are used to produce hosts cells which express the
XX CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX CC used to vaccinate subjects and to raise antibodies against the bacteria.

CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 260 AA;

Query Match 64.9%; Score 943; DB 22; Length 260;
Best Local Similarity 71.3%; Pred. No. 2.3e-78;
Matches 181; Conservative 36; Mismatches 37; Indels 0; Gaps 0;
34 MLYSLNAGGKRRPVLVLLTLDLSLNTYELGKMSAIALEMTHTYSLIHDDLPAMDNDYR 93
1 MLYSLNAGGKRRPVLVLLTLDLSLNTYELGKMSAIALEMTHTYSLIHDDLPAMDNDYR 60
94 RGKLTNKHVYGEWTAIAGDALLKAFELISSDDRLTDEVKIKVLORLSIASGHVGVGG 153
61 RGKLTNKHVYGEWTAIAGDALLKAFELISSDDRLTDEVKIKVLORLSIASGHVGVGG 120
154 QMLDMQSGEQPIDLEMTIHKTKTGALLTFVAVMSAADIANVDDTTKEHLESYSYHLGMM 213
121 QALDMESEKSGIRLETLESIHETKGTALLNFSVAADVIAQVEQNAKNLDEFSHLGMM 180
214 FQIKDDLLDCYGDGAKGKGVGSLNNKSTYVSLGKGADKLTYYHRDAVDELTD 273
181 FQIKDDLLDYVGDSEKLGKGVSDIVNKHSTYVSLGKGAEKLNHNQYLAAMNCLNQS 240
274 EQFNTKHLLEIVDL 287
241 DQYDTSELSDIVDL 254

RESULT 7
ABB48306
ID ABB48306 standard; Protein; 293 AA.

XX ABB48306;

DT 05-FEB-2002 (first entry)

XX Listeria monocytogenes protein #1010.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

PN WO200177335-A2.

XX 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR01118.

XX 11-APR-2000; 2000FR-0004629.

XX (INSP) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kretz J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durand L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;

XX WPI; 2002-010914/01.
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides
XX Claim 6; SEQ ID No 1011; 192pp; French.

XX The present invention relates to the genome sequence of Listeria
CC monocytogenes EGG-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present invention is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC antibodies, identification of L. monocytogenes and related organisms,
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 293 AA;

Query Match 48.3%; Score 702.5; DB 23; Length 293;
Best Local Similarity 52.9%; Pred. No. 3.1e-56;
Matches 148; Conservative 46; Mismatches 81; Indels 5; Gaps 2;

QY 8 KLIDVNNELSVAINKSVMDTOLESMLYSLNAGGKRRPVLVLLTLDLSLNTYELGKMS 67
DB 13 KVIDE---SLPKEINERIEPRKESMLYSIQAGKRRPMLVFATLQALKVNPILGVKT 59
QY 68 ATALEMHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALLKAFELISSDD 127
DB 70 ATALEMHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALLKAFELISSDD 129
QY 128 RLTDEVKIKVLORLSIASGHVGVGQMLDMQSEQPIDLEMTIHKTKTALLTFVAVM 187
DB 130 NLSFETRIALINQISFSSGAEQMGVGGQLADLEAKNKQVTLLELSIHARKTKGELLYAVT 189
QY 188 SAADIANVDDTTKEHLESYSYHLGMMFOIKDOLLDCYGDGAKLGGKVGSDLENNKSTYVS 247
DB 190 SAAKTAADPEQTKRLRIFAENIGIFQISDDILDVIGDETGMKGTGADFLNKSTYPG 249
QY 248 LIGKPDGAEKLTYYHRDAVDELTDIQIDQFNTKHLLEIVDL 287
DB 250 LITLDGAKRALNEHVHTIAKSALSAGHD--FDDEILLKLADL 287

RESULT 8

AAW00285

XX AAW00285 standard; Protein; 297 AA.

AC AAW00285;

DT 11-MAY-1997 (first entry)

XX Mutant farnesylidiphosphate synthase (4).

XX Farnesylidiphosphate: fpp; synthase; mutant; enzyme;
KW geranylgeranyl diphosphate; GGPP.

OS Bacillus stearothermophilus.

XX EP733709-A2.

PN 25-SEP-1996.

```
XX 29-SEP-1995; 95EP-0115423.
XX 14-FEB-1995; 95JP-0025253.
XX (TOYT ) TOYOTA JIDOSHA KK.
XX Ayumi K, Kyoze O, Shinichi O, Shusel O, Takeshi N;
XX Tanetoshi K, Tokuzo N;
XX WPI: 1996-427057/43.
XX N-PSDB; AAT40228.
XX Production of geranyl:geranyl di:phosphate synthase enzymes - by
XX mutation of farnesyl di:phosphate synthase
XX Claim 12; Page 15-16; 50pp; English.
XX Four mutant FPP synthase genes are given in AAT40225 to AAT40228. The
XX native FPP synthase gene is given in AAT40229. The mutants are
XX capable of synthesising geranylgeranyl diphosphate (GGPP) or
XX geranylgeranyl from isopentenyl, dimethylallyl, geranyl or farnesyl
XX diphosphate. The GGPP synthase is produced from FPP synthase by
XX deletion, addition or replacement of one to a few amino acids. The
XX modification is present in at least one of the following positions:
XX 34, 59, 81, 157, 182, 239, 265 and 275.
XX SQ Sequence 297 AA;
Query Match 45.2%; Score 657; DB 17; Length 297;
Best Local Similarity 47.7%; Pred. NO. 4.7e-52;
Matches 133; Conservative 57; Mismatches 83; Indels 6; Gaps 2;
QY 1 MTNLPNKLIDEVNNELSVAINKSVN----DTOLESMYSLNAGGKRIRPVLLLTLDLS 56
DB 1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPALKKAMAYSLEAGGKRIRPLLLSTVRA 60
QY 57 LNTYEELGKMSATALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALL 116
DB 61 LGKDPVAGLPVACAIEMIHTYSLIHDDLPAMDNDLRRGKPTNKHVYGEWTAIAGDGLL 120
QY 117 TKAFELIS--SDDLRTDEVKIKVLQRLSTASGHVGVGOMLDMQSEGQPIDLETLEMIH 174
DB 121 TYAFQITEIDDERIPPSVRLRIERLAKAAGPEGWVAGQAADMEGEGKTLTLESELEYIH 180
QY 175 KTKTGALLTFVNSAADIANDVDTTTRHELESYSYHLGMMFQIKDDLLDCYGDGKAKGKV 234
DB 181 RHKTGKMLQYSVHAGALIGGADARQTRDELDEFAHLGLAFQIRDDILDIEGAEEKIGKPV 240
235 GSDLENNKSTYVSLGKGAEDKLTYHRDAAVDELFTQID 273
DB 241 GSDQSNKATYPALLSLAGAKKELTFHIEAAQRHLRNAD 279
RESULT 9
AAW00286
ID AAW00286 standard; Protein; 297 AA.
XX AC AAW00286;
XX AC AAW00286;
XX AC AAW00286;
DT 11-MAY-1997 (first entry)
DE Native farnesyl diphosphate synthase.
XX Farnesyl diphosphate; FPP; synthase; mutant; enzyme;
XX geranylgeranyl diphosphate; GGPP.
XX Bacillus stearothermophilus.
XX EF733709-A2.
XX 25-SEP-1996.
XX
```

```
PF 29-SEP-1995; 95EP-0115423.
XX 14-FEB-1995; 95JP-0025253.
XX (TOYT ) TOYOTA JIDOSHA KK.
XX Ayumi K, Kyoze O, Shinichi O, Shusel O, Takeshi N;
XX Tanetoshi K, Tokuzo N;
XX WPI: 1996-427057/43.
XX N-PSDB; AAT40229.
XX Production of geranyl:geranyl di:phosphate synthase enzymes - by
XX mutation of farnesyl di:phosphate synthase
XX Disclosure; Page 17-19; 50pp; English.
XX Four mutant FPP synthase genes are given in AAT40225 to AAT40228. The
XX native FPP synthase gene is given in AAT40229. The mutants are
XX capable of synthesising geranylgeranyl diphosphate (GGPP) or
XX geranylgeranyl from isopentenyl, dimethylallyl, geranyl or farnesyl
XX diphosphate. The GGPP synthase is produced from FPP synthase by
XX deletion, addition or replacement of one to a few amino acids. The
XX modification is present in at least one of the following positions:
XX 34, 59, 81, 157, 182, 239, 265 and 275.
XX SQ Sequence 297 AA;
Query Match 45.1%; Score 656; DB 17; Length 297;
Best Local Similarity 47.7%; Pred. NO. 5.8e-52;
Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;
QY 1 MTNLPNKLIDEVNNELSVAINKSVN----DTOLESMYSLNAGGKRIRPVLLLTLDLS 56
DB 1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPALKKAMAYSLEAGGKRIRPLLLSTVQA 60
QY 57 LNTYEELGKMSATALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALL 116
DB 61 LGKDPVAGLPVACAIEMIHTYSLIHDDLPAMDNDLRRGKPTNKHVYGEWTAIAGDGLL 120
QY 117 TKAFELIS--SDDLRTDEVKIKVLQRLSTASGHVGVGOMLDMQSEGQPIDLETLEMIH 174
DB 121 TYAFQITEIDDERIPPSVRLRIERLAKAAGPEGWVAGQAADMEGEGKTLTLESELEYIH 180
QY 175 KTKTGALLTFVNSAADIANDVDTTTRHELESYSYHLGMMFQIKDDLLDCYGDGKAKGKV 234
DB 181 RHKTGKMLQYSVHAGALIGGADARQTRDELDEFAHLGLAFQIRDDILDIEGAEEKIGKPV 240
235 GSDLENNKSTYVSLGKGAEDKLTYHRDAAVDELFTQID 273
DB 241 GSDQSNKATYPALLSLAGAKKELTFHIEAAQRHLRNAD 279
RESULT 10
AAW62532
ID AAW62532 standard; Protein; 297 AA.
XX AC AAW62532;
XX AC AAW62532;
XX AC AAW62532;
DT 14-SEP-1998 (first entry)
DE Farnesyl diphosphate synthase of B. stearothermophilus.
XX Farnesyl diphosphate synthase; enzyme; prenyl diphosphate synthase;
XX PDPS enzyme.
XX Bacillus stearothermophilus.
XX Key Location/Qualifiers
XX FT Misc-difference 1
XX FT /note- "encoded by GTG"
XX FT 86..92
XX FT Domain
XX FT /note- "Asp-rich domain"
```


PN EP733709-A2.
XX 25-SEP-1996.
XX 29-SEP-1995; 95EP-0115423.
XX 14-FEB-1995; 95JP-0025253.
XX (TOYT) TOYOTA JIDOSHA KK.
XX Ayumi K, Kyoze O, Shinichi O, Shusei O, Takeshi N;
PI Tanetoshi K, Tokuzo N;
XX WPI: 1996-427057/43.
DR N-PSDB; AAT40226.
XX Production of geranyl:geranyl di:phosphate synthase enzymes - by
PT mutation of farnesyl di:phosphate synthase
XX Claim 10; Page 10-12; 50pp; English.
XX Four mutant FPP synthase genes are given in AAT40225 to AAT40228. The
CC native FPP synthase gene is given in AAT40229. The mutants are
CC capable of synthesising geranylgeranyl diphosphate (GGPP) or farnesyl
CC geranylgeranyl from isopentenyl, dimethylallyl, geranyl or farnesyl
CC diphosphate. The GGPP synthase is produced from FPP synthase by
CC deletion, addition or replacement of one to a few amino acids. The
CC modification is present in at least one of the following positions:
CC 34, 59, 81, 157, 182, 239, 265 and 275.
XX SQ Sequence 297 AA;
Query Match 44.9%; Score 653; DB 17; Length 297;
Best Local Similarity 47.3%; Pred. No. 1.1e-51;
Matches 132; Conservative 56; Mismatches 85; Indels 6; Gaps 2;
QY 1 MTNLPNKLIDEVNNELSVAINKSVM-----DTQLEESMLYSLNAGGKRIRPVLLLTLDLS 56
DB 1 MAQLSVQEFLEQKQAVETALSRIERLEGPAKVKKAMAYSLEAGGKRIRPLLLSTVQA 60
QY 57 LNTEYELGKMSALALEMIHTYSLIHDDLPAMDNDYRRGKLTNNKVVYGEWTAILAGDALL 116
DB 61 LGDPAVGLPVACAIEMIHTYSLIHDDLPAMDNDYRRGKLTNNKVVYGEWTAILAGDALL 120
QY 117 TFAFELIS--SDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
DB 121 TYAFQLITEIDDERIPPSVRLRIERLAKAAGPEGMAAGQADMEGEGKTLTLESELEYIH 180
175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLLDCYDGAELGKKV 234
DB 181 RHKTGKMLQSVHAGALIGGADARQTRDELDEFAAHLGLAFQIRDDILDIEGAEEKIKPV 240
QY 235 GSDLENNKSYVSLGKGAEKLTVHRDAADVDELTD 273
DB 241 GSDQSNKATYPALLSLAGAKELAFHIEAQAQRHLNAD 279
RESULT 13
AAW00284
ID AAW00284 standard; Protein; 297 AA.
XX AAW00284;
AC AAW00284;
XX 11-MAY-1997 (first entry)
XX Mutant farnesyl diphosphate synthase (3).
DE Farnesyl diphosphate; FPP; synthase; mutant; enzyme;
KW geranylgeranyl diphosphate; GGPP.
XX Bacillus stearothermophilus.
OS Bacillus stearothermophilus.
XX EP733709-A2.

XX 25-SEP-1996.
XX 29-SEP-1995; 95EP-0115423.
XX 14-FEB-1995; 95JP-0025253.
XX (TOYT) TOYOTA JIDOSHA KK.
XX Ayumi K, Kyoze O, Shinichi O, Shusei O, Takeshi N;
PI Tanetoshi K, Tokuzo N;
XX WPI: 1996-427057/43.
DR N-PSDB; AAT40227.
XX Production of geranyl:geranyl di:phosphate synthase enzymes - by
PT mutation of farnesyl di:phosphate synthase
XX Claim 11; Page 13-14; 50pp; English.
XX Four mutant FPP synthase genes are given in AAT40225 to AAT40228. The
CC native FPP synthase gene is given in AAT40229. The mutants are
CC capable of synthesising geranylgeranyl diphosphate (GGPP) or farnesyl
CC geranylgeranyl from isopentenyl, dimethylallyl, geranyl or farnesyl
CC diphosphate. The GGPP synthase is produced from FPP synthase by
CC deletion, addition or replacement of one to a few amino acids. The
CC modification is present in at least one of the following positions:
CC 34, 59, 81, 157, 182, 239, 265 and 275.
XX SQ Sequence 297 AA;
Query Match 44.7%; Score 650; DB 17; Length 297;
Best Local Similarity 47.3%; Pred. No. 2e-51;
Matches 132; Conservative 55; Mismatches 86; Indels 6; Gaps 2;
QY 1 MTNLPNKLIDEVNNELSVAINKSVM-----DTQLEESMLYSLNAGGKRIRPVLLLTLDLS 56
DB 1 MAQLSVQEFLEQKQAVETALSRIERLEGPAKVKKAMAYSLEAGGKRIRPLLLSTVRA 60
QY 57 LNTEYELGKMSALALEMIHTYSLIHDDLPAMDNDYRRGKLTNNKVVYGEWTAILAGDALL 116
DB 61 LGDPAVGLPVACAIEMIHTYSLIHDDLPAMDNDYRRGKLTNNKVVYGEWTAILAGDALL 120
QY 117 TFAFELIS--SDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
DB 121 TYAFQLITEIDDERIPPSVRLRIERLAKAAGPEGMAAGQADMEGEGKTLTLESELEYIH 180
175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLLDCYDGAELGKKV 234
DB 181 RYKTGKMLQSVHAGALIGGADARQTRDELDEFAAHLGLAFQIRDDILDIEGAEEKIKPV 240
QY 235 GSDLENNKSYVSLGKGAEKLTVHRDAADVDELTD 273
DB 241 GSDQSNKATYPALLSLAGAKELAFHIEAQAQRHLNAD 279
RESULT 14
AAW47444
ID AAW47444 standard; Protein; 297 AA.
XX AAW47444;
AC AAW47444;
XX 05-JUN-1998 (first entry)
XX Bacillus stearothermophilus farnesyl diphosphate synthase.
DE Farnesyl diphosphate synthase; prenyl diphosphate synthase; mutant;
KW prenyl diphosphate.
XX Bacillus stearothermophilus.
OS Bacillus stearothermophilus.
XX EP816490-A2.

PD 07-JAN-1998.

XX 02-JUL-1997; 97EP-0111022.

XX 03-JUL-1996; 96JP-0191635.

XX (TOYT) TOYOTA JIDOSHA KK.

XX Ishida C, Narita K, Nishino T, Ohnuma S, Ohto C;

PI Takeuchi Y;

XX WPI; 1998-054911/06.

XX N-PSDB; AAV18516.

PT Mutant prenyl di:phosphate synthase - useful for producing

PT long-chain prenyl di:phosphate

PS Claim 4; Page -: 25pp; English.

XX The present sequence is Bacillus stearothermophilus farnesyl

XX diphosphate synthase, an example of a prenyl diphosphate synthase

XX (PDS).

XX A mutant PDS comprises an amino acid substitution located at the

XX 5th position in the amino-terminal direction from the 1st Asp of

XX the N-terminal of the Asp-rich domain Asp-Asp-Xaa-Xaa-(Xaa-Xaa)-Asp

XX (the Xaa in the parentheses may be absent), which is present in the

XX 2nd region among the conserved regions of wild type PDS. The mutant

XX PDS can be used to produce a prenyl diphosphate having at least 20

XX carbons from an isopentenyl diphosphate, dimethylallyl diphosphate,

XX geranyl diphosphate, farnesyl diphosphate or geranylgeranyl

XX diphosphate substrate. The substituted amino acid is involved in

XX controlling the chain length of the prenyl diphosphate produced.

XX N.B. Sequence not given in the specification, but decoded from the

XX DNA sequence AAV18516.

XX

SQ Sequence 297 AA;

Query Match 44.7%; Score 650; DB 19; Length 297;

Best Local Similarity 49.6%; Pred. No. 2e-51;

Matches 134; Conservative 49; Mismatches 85; Indels 2; Gaps 1;

QY 6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDLSLNTYEELGM 65

DB 10 LNEQKQAVETALSRYIERLEGPAKLKAMAYSLGAGGKRIRPVLLLLTLDLSLNTYEELGM 69

QY 66 KSAIALEMTHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGMQSEGQPIDLETLEMIHKTGTGALLT 124

DB 70 PVACALEMTHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGMQSEGQPIDLETLEMIHKTGTGALLT 129

DB 125 -SDDRLTDEVKIKVQLRSLIAGSHGVGMQSEGQPIDLETLEMIHKTGTGALLT 183

DB 130 IDDERIPPSVRLRLIERLAKAAGPEGWAGVAGAAADMEGEGKTLTLESELEYIHRHKTGKMLQ 189

QY 184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLDCYGDGKAGKGVGSDLENKNS 243

DB 190 YSVHAGALIGGADARQTRDELDEFAHGLGAFQIRDDILDIEGAEKIGKVPVGSQSNKA 249

QY 244 TVVSLGKDGAEKLTTHRDAAVDELTD 273

DB 250 TYPALLSLAGAKEKLAFFHIEAQRHLNAD 279

RESULT 15

AAV27008

ID AAY27008 standard; Protein; 297 AA.

XX AAY27008;

XX 20-SEP-1999 (first entry)

XX Farnesyl diphosphate synthase.

XX Geranyl diphosphate synthase; farnesyl diphosphate synthase; enzyme;

KW

KW

XX

OS

XX

PN

XX

XX

PD

XX

XX

PF

XX

PR

XX

XX

PA

XX

XX

PI

XX

XX

DR

XX

DR

XX

XX

PT

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

XX

XX

SQ

Bacillus stearothermophilus; fermentation; geranyl monoterpene.

Bacillus stearothermophilus.

WO9931254-A1.

24-JUN-1999.

10-DEC-1998; 98WO-JP05590.

16-DEC-1997; 97JP-0346686.

(TOYT) TOYOTA JIDOSHA KK.

Narita K, Nishino T, Ohnuma S, Ohto C;

WPI; 1999-430040/36.

N-PSDB; AAX86781.

Geranyl diphosphate synthase and gene encoding it

Disclosure; Page 35-37; 41pp; Japanese.

The invention provides a geranyl diphosphate synthase which is derived from the farnesyl diphosphate synthase of *Bacillus stearothermophilus* by replacement of the serine residue at position 82 with another residue of higher molecular weight. The enzyme is used in the fermentation and enzyme industries for efficient production of geranyl monoterpenes. The present sequence represents a farnesyl diphosphate synthase.

Sequence 297 AA;

Query Match 44.7%; Score 650; DB 20; Length 297;

Best Local Similarity 49.6%; Pred. No. 2e-51;

Matches 134; Conservative 49; Mismatches 85; Indels 2; Gaps 1;

QY 6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDLSLNTYEELGM 65

DB 10 LNEQKQAVETALSRYIERLEGPAKLKAMAYSLGAGGKRIRPVLLLLTLDLSLNTYEELGM 69

QY 66 KSAIALEMTHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGMQSEGQPIDLETLEMIHKTGTGALLT 124

DB 70 PVACALEMTHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGMQSEGQPIDLETLEMIHKTGTGALLT 129

QY 125 -SDDRLTDEVKIKVQLRSLIAGSHGVGMQSEGQPIDLETLEMIHKTGTGALLT 183

DB 130 IDDERIPPSVRLRLIERLAKAAGPEGWAGVAGAAADMEGEGKTLTLESELEYIHRHKTGKMLQ 189

QY 184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLDCYGDGKAGKGVGSDLENKNS 243

DB 190 YSVHAGALIGGADARQTRDELDEFAHGLGAFQIRDDILDIEGAEKIGKVPVGSQSNKA 249

QY 244 TVVSLGKDGAEKLTTHRDAAVDELTD 273

DB 250 TYPALLSLAGAKEKLAFFHIEAQRHLNAD 279

Search completed: May 29, 2003, 11:58:54

Job time : 36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:58:20 ; Search time 15 Seconds
(without alignments)
562.959 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPNKLIDEVNNELSV.....ELTQIDQFNKHLLEIVDL 287

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1436	98.8	293	3	US-09-276-873-2
2	1015	69.9	302	4	US-09-134-001C-3171
3	657	45.2	297	1	US-08-534-910B-9
4	654	45.0	297	3	US-08-886-466-2
5	654	45.0	297	4	US-09-475-304-2
6	654	45.0	297	4	US-09-101-126-3
7	653	44.9	297	1	US-08-534-910B-7
8	650	44.7	297	1	US-08-534-910B-8
9	650	44.7	297	1	US-08-534-910B-10
10	650	44.7	297	4	US-09-367-528A-5
11	646	44.5	297	4	US-09-367-528A-1
12	645	44.4	297	1	US-08-534-910B-6
13	644	44.3	297	4	US-09-367-528A-3
14	551	37.9	291	4	US-09-275-742-2
15	521	35.9	393	3	US-09-187-050-34
16	520.5	35.8	393	3	US-09-187-050-2
17	520.5	35.8	393	3	US-09-187-050-14
18	520.5	35.8	393	3	US-09-187-050-16
19	520.5	35.8	393	3	US-09-187-050-18
20	520.5	35.8	393	3	US-09-187-050-20
21	520.5	35.8	393	3	US-09-187-050-22
22	520.5	35.8	393	3	US-09-187-050-24
23	520.5	35.8	393	3	US-09-187-050-26
24	520.5	35.8	393	3	US-09-187-050-27
25	520.5	35.8	393	3	US-09-187-050-28
26	520.5	35.8	393	3	US-09-187-050-29
27	520.5	35.8	393	3	US-09-187-050-30

28	520.5	35.8	393	3	US-09-187-050-31	Sequence 31, Appl
29	520.5	35.8	393	3	US-09-187-050-32	Sequence 32, Appl
30	520.5	35.8	393	3	US-09-187-050-33	Sequence 33, Appl
31	517	35.6	285	3	US-09-187-050-12	Sequence 12, Appl
32	465.5	32.0	377	4	US-09-420-211-2	Sequence 2, Appl
33	376.5	25.9	302	1	US-07-783-705A-1	Sequence 1, Appl
34	346	23.8	298	1	US-08-095-726-4	Sequence 4, Appl
35	346	23.8	298	1	US-08-096-043-4	Sequence 4, Appl
36	346	23.8	298	1	US-08-093-577-4	Sequence 4, Appl
37	346	23.8	298	1	US-08-096-623A-4	Sequence 4, Appl
38	346	23.8	307	1	US-08-095-726-2	Sequence 2, Appl
39	346	23.8	307	1	US-08-096-043-2	Sequence 2, Appl
40	346	23.8	307	1	US-08-093-577-2	Sequence 2, Appl
41	346	23.8	307	1	US-08-096-623A-2	Sequence 2, Appl
42	330	22.7	330	1	US-08-410-167A-4	Sequence 4, Appl
43	330	22.7	330	2	US-08-898-560-1	Sequence 1, Appl
44	323	22.2	330	4	US-09-101-126-1	Sequence 1, Appl
45	309.5	21.3	295	3	US-08-660-645A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-276-873-2
; Sequence 2, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: ISPA
; FILE REFERENCE: GM10208
; CURRENT APPLICATION NUMBER: US/09/276,873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-276-873-2

Query Match	98.8%	Score 1436;	DB 3;	Length 293;
Best Local Similarity	99.0%	Pred. No. 1e-132;	2;	Indels 0;
Mismatches	284;	Conservative	1;	Mismatches 0;
Gaps	0;			
QY	1	MTNLPNKLIDEVNNELSVAINKSVMDTOLESMLYSINAGGKRIRPVLLLTLDLSLNT	60	
DB	1	MTNLPNKLIDEVNNELSVAINKSVMDTOLESMLYSINAGGKRIRPVLLLTLDLSLNT	60	
QY	61	YELGKMSATALEMIHTYSLIHDDLPAONDYRRGKLTNKHVYGEWTAIAGDALTKAF	120	
DB	61	YELGVKSAIALEMIHTYSLIHDDLPAONDYRRRRLTNKHVYGEWTAIAGDALTKAF	120	
QY	121	ELISSDDRDLTVKTKVLRSLTASCHVGVGQMLDMQSEGPIDLETFLEMIHTKTCA	180	
DB	121	ELISSDDRDLTVKTKVLRSLTASCHVGVGQMLDMQSEGPIDLETFLEMIHTKTCA	180	
QY	181	LITFAVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLLDCYGEAKLKKVGSLEN	240	
DB	181	LITFAVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLLDCYGEAKLKKVGSLEN	240	
QY	241	NKSTVSLGKGAEDKLTTHRDAAYDELTDQFNTHKHLLEIVDL	287	
DB	241	NKSTVSLGKGAEDKLTTHRDAAYDELTDQFNTHKHLLEIVDL	287	

RESULT 2
US-09-134-001C-3171
; Sequence 3171, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GFC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3171
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3171

Query Match 69.9%; Score 1015; DB 4; Length 302;
Best Local Similarity 68.6%; Pred. No. 1.7e-91;
Matches 197; Conservative 41; Mismatches 49; Indels 0; Gaps 0;

Db 1 MTNLPNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLILLTSLNTE 60
10 MKKLQMNKLIINTSLNKSIQSPKLTNLEESMKYSLNAGGRIRPVLILLTLKMLNKD 69

QY 61 YELGMSAIALEMIHYSLIHDDLPAMDNDYRRGKLTNHHKYVGEWTAIAGDALLTKAF 120
70 YQGLNSALALEMIHYSLIHDDLPAMDNDYRRGKLTNHHKYVGEWTAIAGDALLTKAF 129

QY 121 ELTSSDRRLTDEVKIKVQLRSLIASGHVGMVGMQMDQSEGOPIDLETLEMTHTKGTGA 180
130 ELVSNDDTTEDSVKVIKRLKASGHLGVGMQMDQSEGOPIDLETLEMTHTKGTGA 189

QY 181 LLTFVMSAADIANVDDTTKEHLESYSHGLMFMFOIKDDLLDCYGDGKLVKGVSDLEN 240
190 LLNFSVMAAVDIAQVEQNIAKNLDEFSHLGMFMFOIKDDLLDYGDGSKLVKGVSDIVN 249

QY 241 NKSTYVSLGKGAEKDLTHRDAADVDELTOIDEQNTKHLLEIVDL 287
250 HKSTYVSLGKGAEKDLTHRDAADVDELTOIDEQNTKHLLEIVDL 296

RESULT 3
US-08-534-910B-9
; Sequence 9, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesylphosphate Synthase Capable
; OF SYNTHESIZING GERANYLERANYLDIPHOSPHATE AND GENE CODING THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.
; ZIP: 20036-3405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" Floppy Disk
; OPERATING SYSTEM: IBM PC Compatible
; SOFTWARE: IBM/Word Perfect 6.1 Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/534,910B
; FILING DATE: 28-SEPT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,001C
; FILING DATE: 14-FEB-1995
; NAME: Toffenetti, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 77670/398
; TELEPHONE: (202)429-1776
; TELEFAX: (202)429-0796
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORGANISM: Bacillus stearothermophilus
; US-08-534-910B-9

Query Match 45.2%; Score 657; DB 1; Length 297;
Best Local Similarity 47.7%; Pred. No. 1.8e-56;
Matches 133; Conservative 57; Mismatches 83; Indels 6; Gaps 2;

QY 1 MTNLPNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLILLTSLNTE 56
1 MAQLSVSEQFLNEQKQAVETALSRYIERLEGPAKLKAMAYSLEAGGRIRPVLILLSTVRA 60

QY 57 LNTPEYELGMSAIALEMIHYSLIHDDLPAMDNDYRRGKLTNHHKYVGEWTAIAGDALL 116
61 LGHDPAVLGPVACAIEMIHYSLIHDDLPAMDNDYRRGKLTNHHKYVGEWTAIAGDALL 120

QY 117 TKAFELIS--SDRLTDEVKIKVQLRSLIASGHVGMVGMQMDQSEGOPIDLETLEMIH 174
121 TYAQLITTEIDDERIPPSVRLRIERLAKAAGPEGVAGQADMEGEGTKLTSLSELYIH 180

QY 175 KTTGALLTFVMSAADIANVDDTTKEHLESYSHGLMFMFOIKDDLLDCYGDGKLVK 234
181 RHKTGKMLQYSHVAGALIGGADARQTRDELDEFAHLGLAFQIRDDILIDIEGAEKIKRY 240

QY 235 GSDLENNKSTYVSLGKGAEKDLTHRDAADVDELTOID 273
241 GSDQSNKATYPALLSLAGAKELTFHIEAQRHLNAD 279

RESULT 4
US-08-886-466-2
; Sequence 2, Application US/0886466C
; Patent No. 6040165
; GENERAL INFORMATION:
; APPLICANT: Narita, Keishi
; APPLICANT: Ishida, Chika
; APPLICANT: Takeuchi, Yoshie
; APPLICANT: Ohto, Chikara
; APPLICANT: Ohnuma, Shinichi
; APPLICANT: Nishino, Tokuzo
; TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
; FILE REFERENCE: 77670/494
; CURRENT APPLICATION NUMBER: US/08/886,466C
; CURRENT FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: JP 8-191635
; EARLIER FILING DATE: 1996-07-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
; US-08-886-466-2

Query Match 45.0%; Score 654; DB 3; Length 297;
Best Local Similarity 47.7%; Pred. No. 3.6e-56;
Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;
```

QY 1 MTNLPNKLIDEVNNELSVAINKSVN----DTQLEESMLYSLNAGGKRIRPVLLLTLD 56
Db 1 MAQLSVEQFLNPKQAVETALSRYTERLEGPAAKLKAMAYSLEAGGKRIRPVLLLTSTVRA 60
QY 57 LNTVEYELGMSAIALEMTHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGTAILAGDALL 116
Db 61 LKDPVAVGLPVACAIEMHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGTAILAGDALL 120
QY 117 TKAFELIS--SDRLTDEVKIKVLQRLSIASGHVGVGGQMDQSEGQPIDLETLEMH 174
Db 121 TYAFQLITEIDDERIPPSVRLRIERLAKAAGPEGWAGVAGQADMEGEGKTLTLESEYIH 180
QY 175 KTKTGALLTFVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLCCYGDGKGLKV 234
Db 181 RHKTGMQLQYSVHAGALIGGADAROTRELDFAAHLGLAFQIRDDILDIEGAEEKIGRPV 240
QY 235 GSDLENNKSTVYSLGKGAEDKLTYYHRDAVDELTD 273
Db 241 GSDQSNKATYPALLSLAGAKEKLAFFHIEAQRHLRNAD 279

RESULT 5

US-09-475-304-2

Sequence 2, Application US/09475304

Patent No. 6225096

GENERAL INFORMATION:

APPLICANT: Narita, Keishi

APPLICANT: Ishida, Chika

APPLICANT: Takeuchi, Yoshie

APPLICANT: Ohto, Chikara

APPLICANT: Ohnuma, Shinichi

APPLICANT: Nishino, Tokuzo

FILE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE

FILE REFERENCE: 77670/494

CURRENT APPLICATION NUMBER: US/09/475,304

EARLIER FILING DATE: 1999-12-30

EARLIER APPLICATION NUMBER: JP 8-191635

EARLIER FILING DATE: 1996-07-03

NUMBER OF SEQ ID NOS: 9

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 297

TYPE: PRT

ORGANISM: Bacillus stearothermophilus

US-09-475-304-2

Query Match 45.0%; Score 654; DB 4; Length 297;

Best Local Similarity 47.7%; Pred. No. 3.6e-56;

Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;

QY 1 MTNLPNKLIDEVNNELSVAINKSVN----DTQLEESMLYSLNAGGKRIRPVLLLTLD 56
Db 1 MAQLSVEQFLNPKQAVETALSRYTERLEGPAAKLKAMAYSLEAGGKRIRPVLLLTSTVRA 60
QY 57 LNTVEYELGMSAIALEMTHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGTAILAGDALL 116
Db 61 LKDPVAVGLPVACAIEMHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGTAILAGDALL 120
QY 117 TKAFELIS--SDRLTDEVKIKVLQRLSIASGHVGVGGQMDQSEGQPIDLETLEMH 174
Db 121 TYAFQLITEIDDERIPPSVRLRIERLAKAAGPEGWAGVAGQADMEGEGKTLTLESEYIH 180
QY 175 KTKTGALLTFVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLCCYGDGKGLKV 234
Db 181 RHKTGMQLQYSVHAGALIGGADAROTRELDFAAHLGLAFQIRDDILDIEGAEEKIGRPV 240
QY 235 GSDLENNKSTVYSLGKGAEDKLTYYHRDAVDELTD 273
Db 241 GSDQSNKATYPALLSLAGAKEKLAFFHIEAQRHLRNAD 279

RESULT 6

US-09-101-126-3

Sequence 3, Application US/09101126
Patent No. 6316216
GENERAL INFORMATION:
APPLICANT: OHTO, CHIKARA
APPLICANT: NAKANE, HIROYUKI
APPLICANT: NISHINO, TOKUZO
APPLICANT: OHNUMA, SHINICHI
APPLICANT: HIROOKA, KAZUTAKE
FILE OF INVENTION: MUTATED PRENYL DIPHOSPHATE SYNTHASE
FILE REFERENCE: 77670/566
CURRENT APPLICATION NUMBER: US/09/101,126
CURRENT FILING DATE: 1999-04-27
EARLIER APPLICATION NUMBER: PCT/JP97/03921
EARLIER FILING DATE: 1997-10-29
EARLIER APPLICATION NUMBER: JP 8-307506
EARLIER FILING DATE: 1996-11-05
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 297
TYPE: PRT
ORGANISM: Bacillus stearothermophilus
FEATURE:
OTHER INFORMATION: 86-92 is an Asp-rich domain
US-09-101-126-3

Query Match 45.0%; Score 654; DB 4; Length 297;

Best Local Similarity 47.7%; Pred. No. 3.6e-56;

Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;

QY 1 MTNLPNKLIDEVNNELSVAINKSVN----DTQLEESMLYSLNAGGKRIRPVLLLTLD 56
Db 1 MAQLSVEQFLNPKQAVETALSRYTERLEGPAAKLKAMAYSLEAGGKRIRPVLLLTSTVRA 60
QY 57 LNTVEYELGMSAIALEMTHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGTAILAGDALL 116
Db 61 LKDPVAVGLPVACAIEMHTYSLIHDDLPAMDNDYRRGKLTNHNKVGWGTAILAGDALL 120
QY 117 TKAFELIS--SDRLTDEVKIKVLQRLSIASGHVGVGGQMDQSEGQPIDLETLEMH 174
Db 121 TYAFQLITEIDDERIPPSVRLRIERLAKAAGPEGWAGVAGQADMEGEGKTLTLESEYIH 180
QY 175 KTKTGALLTFVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLCCYGDGKGLKV 234
Db 181 RHKTGMQLQYSVHAGALIGGADAROTRELDFAAHLGLAFQIRDDILDIEGAEEKIGRPV 240
QY 235 GSDLENNKSTVYSLGKGAEDKLTYYHRDAVDELTD 273
Db 241 GSDQSNKATYPALLSLAGAKEKLAFFHIEAQRHLRNAD 279

RESULT 7

US-08-534-910B-7

Sequence 7, Application US/08534910B

Patent No. 5766911

GENERAL INFORMATION:

APPLICANT: KOIKE, Ayumi

APPLICANT: OBATA, Shusei

APPLICANT: NISHINO, Tokuzo

APPLICANT: OHNUMA, Shinichi

APPLICANT: NAKAZAWA, Takeshi

APPLICANT: OGURA, Kyoza

APPLICANT: KOYAMA, Tanetoshi

FILE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable

TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: 1025 Connecticut Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: U.S.

ZIP: 20036-5405

```
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.25" Floppy Disk
;; COMPUTER: IBM PC Compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
;; SOFTWARE: IBM/Word Perfect 6.1 Windows
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/534,910B
;; FILING DATE: 28-SEPT-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 7-25253
;; FILING DATE: 14-FEB-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Toffenetti, Judith L.
;; REGISTRATION NUMBER: 39,048
;; REFERENCE/DOCKET NUMBER: 77670/398
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)429-1776
;; TELEFAX: (202)429-0796
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 297 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Bacillus stearothermophilus
;; US-08-534-910B-7

Query Match 44.98; Score 653; DB 1; Length 297;
Best Local Similarity 47.38; Pred. No. 4.5e-56;
Matches 132; Conservative 56; Mismatches 85; Indels 6; Gaps 2;

QY 1 MTNLPNKLDEYNNELSVAINKSVM----DTOLEESMLYSLNAGGKRIRPVLILLTLDLS 56
Db 1 MAQLSVQEFLENEQAVETALSRYIERLEGPAKVKAMAYSLEAGGKRIRPVLILLTLDLS 60
QY 57 LNTYEYELGKMSATALEMIHTYSLIHDDLPAMDNDYRRGKLTNKKYVGEWTAIAGDALL 116
Db 61 LGKDPVAVGLPVACAIEMIHTYSLIHDDLPAMDNDYRRGKLTNKKYVGEWTAIAGDALL 120
QY 117 TKAFELIS--SDRLTDEVKIKVLRSLTASGHVGVGQMDQSGOPIDLETLEMIH 174
Db 121 TYAQLITEIDDERIPPSVRLRIERLAKAAGPEGMAAGQADMEGEGKTLTLESELEYIH 180
QY 175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMFMFOIKDDLLDCYGDRAKLGKV 234
Db 181 RHTGKMLQSVHAGALIGGADARQTRDELDEFAHGLGLAFQIRDDIILDIEGAEKIGKPV 240
QY 235 GSDLENNKSTYVSLGKDGAEKLTVHRDAVDELTD 273
Db 241 GSDQSNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279

RESULT 8
US-08-534-910B-8
; Sequence 8, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesylphosphate Synthase Capable
; TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Therefor
; NUMBER OF SEQUENCES: 10
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
```

```
;; STATE: DC
;; COUNTRY: U.S.
;; ZIP: 20036-5405
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.25" Floppy Disk
;; COMPUTER: IBM PC Compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
;; SOFTWARE: IBM/Word Perfect 6.1 Windows
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/534,910B
;; FILING DATE: 28-SEPT-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 7-25253
;; FILING DATE: 14-FEB-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Toffenetti, Judith L.
;; REGISTRATION NUMBER: 39,048
;; REFERENCE/DOCKET NUMBER: 77670/398
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)429-1776
;; TELEFAX: (202)429-0796
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 297 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Bacillus stearothermophilus
;; US-08-534-910B-8

Query Match 44.7%; Score 650; DB 1; Length 297;
Best Local Similarity 47.3%; Pred. No. 8.9e-56;
Matches 132; Conservative 55; Mismatches 86; Indels 6; Gaps 2;

QY 1 MTNLPNKLDEYNNELSVAINKSVM----DTOLEESMLYSLNAGGKRIRPVLILLTLDLS 56
Db 1 MAQLSVQEFLENEQAVETALSRYIERLEGPAKVKAMAYSLEAGGKRIRPVLILLTLDLS 60
QY 57 LNTYEYELGKMSATALEMIHTYSLIHDDLPAMDNDYRRGKLTNKKYVGEWTAIAGDALL 116
Db 61 LGKDPVAVGLPVACAIEMIHTYSLIHDDLPAMDNDYRRGKLTNKKYVGEWTAIAGDALL 120
QY 117 TKAFELIS--SDRLTDEVKIKVLRSLTASGHVGVGQMDQSGOPIDLETLEMIH 174
Db 121 TYAQLITEIDDERIPPSVRLRIERLAKAAGPEGMAAGQADMEGEGKTLTLESELEYIH 180
QY 175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMFMFOIKDDLLDCYGDRAKLGKV 234
Db 181 RYKTGKMLQSVHAGALIGGADARQTRDELDEFAHGLGLAFQIRDDIILDIEGAEKIGKPV 240
QY 235 GSDLENNKSTYVSLGKDGAEKLTVHRDAVDELTD 273
Db 241 GSDQSNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279

RESULT 9
US-08-534-910B-10
; Sequence 10, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesylphosphate Synthase Capable
; TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
```


Query Match	44.5%	Score	646;	DB	4;	Length	297;
Best Local Similarity	49.3%	Pred.	No. 2.2e-55;				
Matches	133;	Conservative	49;	Mismatches	86;	Indels	2;
						Gaps	1;
QY	6	MNKLIDVNNELSVAINKSWDQLESMYLSLNAGGKRIRPVLLLTLDLSLNPEYELGM	65				
DB	10	LNEQKQAVETALSRYIERLEGPALKKAMAYSLEAGGKRIRPLLLSTVRALGRKDPAVGL	69				
QY	66	KSAALEMIITYSLIHDDLPAMDNDYRRGKLTNHHKVYGEWTAIACDALLTKAFELIS-	124				
DB	70	PVACAIEMIHTYXLIHDDLPAMDNDLRRGKPTNHHKVGEAMAIACDGLLTVAFQLITE	129				
QY	125	-SDRLTDEYKIVQLRLSTASGHVGVGQMLDMQSEGQPIDLETLEMTHKTKTGALLT	183				
DB	130	IDDERIPPSVRLRIERLAKAAGPEGVAVAGAADMEGEGTKTLTLESELEYTHRRKTKGMLQ	189				

QY 184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDCAKLGKVGSDLENKKS 243
Db 190 YSVHAGALIGGADARQTRDELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPVGSDQSNKA 249
QY 244 TYVSLGKDGAEKLTYYHRDAAVDELTD 273
Db 250 TYPALLSLAGAKEKLAFFHIEAQRHLRNAD 279

RESULT 12
US-08-534-910B-6
; Sequence 6, Application US/08534910B
; Patent No. 5768911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOTAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesyl diphosphate Synthase Capable
; TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: IBM/Word Perfect 6.1 Windows
; CURRENT APPLICATION DATA: US/08/534,910B
; APPLICATION NUMBER: 7-25253
; FILING DATE: 28-SEPT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE: 14-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Toffenetti, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 77670/398
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)429-1776
; TELEFAX: (202)429-0796
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
US-08-534-910B-6

Query Match 44.4%; Score 645; DB 1; Length 297;
Best Local Similarity 48.0%; Pred. No. 2.8e-55;
Matches 130; Conservative 56; Mismatches 79; Indels 6; Gaps 2;

QY 1 MTNLPNKLTIDEVNNELSVAINKSVN----DTQLESMLYSINAGGKRIRPVLILLTLD 56
Db 1 MAQLSEVQFLNEQKQAVETALSRYTERLEGPALKKAMAYSLEAGGKRIRPILLSTVRA 60
QY 57 LNTVEYLGKMSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALL 116
Db 61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDGL 120
QY 117 TRAFELIS--SDDLRTDEVKIKVQLRSLASGHVGVGMQDLMQSEGQPIDLETLEMIH 174

Db 121 TYAFOLITEIDDRIPPSVRLRIERLAKAAGPEGVAGQAADMEGEGKTLTISELEYIH 180
QY 175 KTKTGALLTFAYMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDCAKLGK 234
Db 181 RHKTGKMLQYSVHAGALIGGADARQTRDELDEFAAHLGLAFQIRDDILDIEGAEEK 240
QY 235 GSDLENKSTYVSLGKDGAEKLTYYHRDAA 265
Db 241 GSDQSNKATYPALLSLAGAKEKLAFFHIEA 271

RESULT 13
US-09-367-528A-3
; Sequence 3, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
; APPLICANT: TOYOTA JIDOSHA KABUSHIKI KAISHA
; TITLE OF INVENTION: Geranyl diphosphate Synthetase Gene
; FILE REFERENCE: PH-586
; CURRENT APPLICATION NUMBER: US/09/367,528A
; CURRENT FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: JP97/346686
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-09-367-528A-3

Query Match 44.3%; Score 644; DB 4; Length 297;
Best Local Similarity 49.3%; Pred. No. 3.5e-55;
Matches 133; Conservative 49; Mismatches 86; Indels 2; Gaps 1;

QY 6 MNKLIDEVNNELSVAINKSVMDTQLESMLYSINAGGKRIRPVLILLTLDLSLNTVEYLG 65
Db 10 LNEQKQAVETALSRYTERLEGPALKKAMAYSLEAGGKRIRPILLSTVRLGKDPVGL 69
QY 66 KSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALLTKAFELIS 124
Db 70 PVACAIEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDGLTYAFOLITE 129
QY 125 -SDDLRTDEVKIKVQLRSLASGHVGVGMQDLMQSEGQPIDLETLEMIHKTGTGALLT 183
Db 130 IDDERIPPSVRLRIERLAKAAGPEGVAGQAADMEGEGKTLTISELEYIHRHKTGKMLQ 189
QY 184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDCAKLGKVGSDLENKKS 243
Db 190 YSVHAGALIGGADARQTRDELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPVGSDQSNKA 249
QY 244 TYVSLGKDGAEKLTYYHRDAAVDELTD 273
Db 250 TYPALLSLAGAKEKLAFFHIEAQRHLRNAD 279

RESULT 14
US-09-275-742-2
; Sequence 2, Application US/09275742
; Patent No. 6130069
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: ispa
; FILE REFERENCE: GM10205
; CURRENT APPLICATION NUMBER: US/09/275,742
; CURRENT FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 291
; TYPE: PRT

ORGANISM: Streptococcus pneumoniae
US-09-273-742-2

Query Match 37.9%; Score 551; DB 4; Length 291;
Best Local Similarity 46.6%; Pred. No. 4.2e-46;
Matches 122; Conservative 48; Mismatches 88; Indels 4; Gaps 3;

QY 28 TQLESMLYSLNAGKRRPVLVLLTLDLSLN-TEYELGKMSAIALEMIHTYSLIHDDLP 86
DB 25 SSLRESVLYSIHAGGRIRPFLLVKLEALQVTKPAHAQVATALEMIHTGSLIHDDLP 84
QY 87 MNDYRRGKLTNKHVGEWTAIAGDALLTKAPELISDDRLTDEVKIKVLRSLASG 146
DB 85 MDDYRRGRLTNHKKFGEAAILAGDALFDYPALIAQAD-LPSQIKVGLIANLSL 143
QY 147 HGVWVGQMLDMSQEGPIDLETLEMIHTKTGALLTFVMSAADIANVDDTTKEHLESY 206
DB 144 SLGMVAGQVLDMEGHEHLSLEELQTIHANKTGKLLAYPFOAAAIIELSPEMQVKLTV 203
QY 207 SYHGMFQIKRDDLLDCYDGAELKGVGSDLENNKSTYVLLGKDGAEKLTYYHRDAV 266
DB 204 GELIGLAFQVRDVLVDVTFASFEICKTPKDLQAEKSTYPALLGLEESIAFCNQTLDQAN 263
QY 267 DELTOIDQO--FNTKHLLEIVD 286
DB 264 DKLEIAAQLPFETESIVSVE 285

RESULT 15

US-09-187-050-34
Sequence 34, Application US/09187050B
Patent No. 6043072

GENERAL INFORMATION:

APPLICANT: Croteau, Rodney B
APPLICANT: Hefner, Jerry
TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranylgeranyl Diphosphate
TITLE OF INVENTION: Synthese, And Methods of Use
FILE REFERENCE: WSUR12423
CURRENT APPLICATION NUMBER: US/09/187,050B
CURRENT FILING DATE: 1998-11-05
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 34
LENGTH: 393

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Taxus GGPP
OTHER INFORMATION: synthase protein variant

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)..(393)

OTHER INFORMATION: Computer-generated Taxus GGPP synthase protein

OTHER INFORMATION: variant

US-09-187-050-34

Query Match 35.9%; Score 521; DB 3; Length 393;
Best Local Similarity 38.3%; Pred. NO. 5.7e-43;
Matches 118; Conservative 55; Mismatches 107; Indels 28; Gaps 5;

QY 6 MNKLLDEVNNE-----LSVAINKSV---MDTQLESMLYSLNAG 41
DB 80 MGQLVDLAETEKAEGKIEFDNFNEYMKS KAVAVDAALDKAIPLEYPERIHESMRYSLLAG 139
QY 42 GKRIRPVLVLLTLDLSNTEYELGKMSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHK 101
DB 140 GKRVPALCIAACEVLVGSQDLAMPFTACAMEMIHTWSLIHDDLPAMDNDYRRGKLTNHK 199
QY 102 YGEWTAIAGDALTKAFE--LISSDDRLTDEVKIKVLRSLASGHVWVGQMLDMQ 159
DB 200 VEGEDTAVLAGDALLSFAFEHIAVATSKTVPSDRLTVISELGKTIQSQGLVGGQVVDIT 259
QY 160 SEGQP-IDLETLEMIHTKTGALLTFVMSAADIANVDDTTKEHLESYSLHGMFQIKND 218

DB 260 SEGDNVDLKTLEWHIHKTAIVLLECSVVSGGILGATEDEIARRRARCVCYGLLFQVVD 319
QY 219 DILLDCYGEAKLGKKVGSLENNKSTYVLLGKDGAEKLTYYHRDAAVDELTDQIDEQFNT 278
DB 320 DILDVTKSSEELGKTAGNDLLTDKATYPKLMGLEKAKEPAEALATPAKEELSSFD-QIKA 378
QY 279 KHLLEIVD 286
DB 379 APPLGLAD 386

Search completed: May 29, 2003, 12:00:37
Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:59:55 ; Search time 45 Seconds
(without alignments)
645.582 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPKMLIDEVNNELSA.....ELTQIDQFNKHLLEIVDL 287

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1453	100.0	287	9	US-10-084-205-64
2	1453	100.0	287	10	US-09-925-637-64
3	1442	99.2	293	10	US-08-815-242-12583
4	1415	97.4	288	10	US-09-815-242-5239
5	596	41.0	293	10	US-08-815-242-10630
6	549	37.8	291	10	US-09-815-242-13597
7	547	37.6	291	10	US-09-815-242-13273
8	516.5	35.5	297	9	US-09-941-947A-20
9	516.5	35.5	297	10	US-09-934-903-14
10	516.5	35.5	297	10	US-09-934-868-72
11	504.5	34.7	367	9	US-10-108-915-26
12	497.5	34.2	295	10	US-09-815-242-11239
13	494.5	34.0	350	9	US-10-108-915-22
14	487.5	33.6	369	9	US-10-108-915-18
15	485.5	33.4	299	10	US-09-815-242-10069
16	478.5	32.9	316	9	US-10-108-915-45
17	474.5	32.7	295	10	US-09-815-242-11971
18	473	32.6	299	10	US-09-815-242-14084
19	465.5	32.0	377	10	US-09-934-778-2

ALIGNMENTS

RESULT 1

US-10-084-205-64
; Sequence 64, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 64
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-084-205-64

Query Match	100.0%	Score 1453;	DB 9;	Length 287;
Best Local Similarity	100.0%	Pred. No. 1.3e-119;		
Matches 287;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTNLPKMLIDEVNNELSAVINKSVMDTQLEESMLYSLNAGGRIRPVLILLTLDLSLNTG	60	
DB	1	MTNLPKMLIDEVNNELSAVINKSVMDTQLEESMLYSLNAGGRIRPVLILLTLDLSLNTG	60	
QY	61	YELGKMSATALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALTKAF	120	
DB	61	YELGKMSATALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALTKAF	120	
QY	121	ELISSDDRLTDEVKTKVLQRLSTASHGVGVGQMLDMQSEGQIDLETLEMIHKTGA	180	
DB	121	ELISSDDRLTDEVKTKVLQRLSTASHGVGVGQMLDMQSEGQIDLETLEMIHKTGA	180	
QY	181	LTFFAVMSAADIANVDDTTKEHLESYSYHLGMVFQIKDLDLCYGDGKLVKVGSDLEN	240	
DB	181	LTFFAVMSAADIANVDDTTKEHLESYSYHLGMVFQIKDLDLCYGDGKLVKVGSDLEN	240	

QY 241 NKSTVSVLLGKDGAEDKLT YHRDAAVDEL TQIDQFNTKHLLEIVDL 287
 |||||
 Db 241 NKSTVSVLLGKDGAEDKLT YHRDAAVDEL TQIDQFNTKHLLEIVDL 287
 |||||

RESULT 2

```

US-09-925-637-64
; Sequence 64, Application US/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 287
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-925-637-64

```

RESULTS

```

RESOLUTION 3
US-09-815-242-12583
; Sequence 12583, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

```

;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5239
;; LENGTH: 288
;; TYPE: PRF
;; ORGANISM: Staphylococcus aureus
US-09-815-242-5239

Query Match 97.4%; Score 1415; DB 10; Length 288;
Best Local Similarity 99.3%; Pred. No. 2.8e-116;
Matches 280; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSINAGGRIRPVLVLLTLDLSLNTYEELGM 65
1 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSINAGGRIRPVLVLLTLDLSLNTYEELGM 60

QY 66 KSAIALEMIHTYSLTHDDLPAMDNDYRRGKLTNHKVTGWTAILAGDALLTKAFELISS 125
61 KSAIALEMIHTYSLTHDDLPAMDNDYRRGKLTNHKVTGWTAILAGDALLTKAFELISS 120

QY 126 DDLRLDEVKIKVQLRLSTAGHVGVMGOMLDMQSEGQPIDLETLEMIHKTGTGALLTFA 185
121 DDLRLDEVKIKVQLRLSTAGHVGVMGOMLDMQSEGQPIDLETLEMIHKTGTGALLTFA 180

QY 186 VMSAADIANDVDTTKEHLESYSYHLGMFQIKDDLLDCYGDGAKLGGKVGSDLENNKSTY 245
181 VMSAADIANDVDTTKEHLESYSYHLGMFQIKDDLLDCYGDGAKLGGKVGSDLENNKSTY 240

QY 246 VSLLGKGAEDKLTTHRDAANDELQIDEQFNTKHLLEIVDL 287
241 VSLLGKGAEDKLTTHRDAANDELQIDEQFNTKHLLEIVDL 282

RESULT 5
US-09-815-242-10630
;; Sequence 10630, Application US/09/815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE OF INVENTION: Prokaryotes
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5239
;; LENGTH: 288
;; TYPE: PRF
;; ORGANISM: Staphylococcus aureus

;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 10630
;; LENGTH: 293
;; TYPE: PRF
;; ORGANISM: Enterococcus faecalis
US-09-815-242-10630

Query Match 41.0%; Score 596; DB 10; Length 293;
Best Local Similarity 46.0%; Pred. No. 1.7e-44;
Matches 134; Conservative 56; Mismatches 91; Indels 10; Gaps 5;

QY 1 MTNLPWN--KLIDEVNNELSVAINKSVMDTQLEESMLYSINAGGRIRPVLVLLTLDLSLN 58
1 MTNFSQOHLPLVEKVMDF--IAEYTERLERKEAMLYSHAGGKRLRLVLTVAAPQ 57

Db 59 TEVEL-GMKSATALEMIHTYSLTHDDLPAMDNDYRRGKLTNHKVTGWTAILAGDALLT 117
58 KEMETQDIQVAAASLEMIHTYSLTHDDLPAMDNDYRRGKLTNHKVTGWTAILAGDALLT 117

QY 118 KAFELISSDDRLTDEVKIKVQLRLSTAGHVGVMGOMLDMQSEGQPIDLETLEMIHKT 177
118 GAFQLLSLQLGSE-KVLLMQQLAKACAGNAGVAGMGDIEGEKYSILTEELAAVHEKK 176

QY 178 TGALLTFAVMSAADIANDVDTTKEHLESYSYHLGMFQIKDDLLDCYGDGAKLGGKVGSD 237
177 TGALLTEFALLTAGVLANQTEEVIGLLTQFAHLYGLAFQIRDDLLDATSTEADLGGKVGSD 236

QY 238 LENNKSTYVSLGKGAEDKLT--HRAANDELQIDEQFNTKHLLEIV 285
237 EALNKSTYPALLGIAGAKDALTQLAEGSAVLEKIKANYPNTSEHLANLL 287

RESULT 6
US-09-815-242-13597
;; Sequence 13597, Application US/09/815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE OF INVENTION: Prokaryotes
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13597
;; LENGTH: 291
;; TYPE: PRF
;; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13597

Query Match 37.8%; Score 549; DB 10; Length 291;
Best Local Similarity 46.2%; Pred. No. 2.3e-40;
Matches 121; Conservative 50; Mismatches 87; Indels 4; Gaps 3;

QY 28 TQLEESMLYSLNAGGKRIRPVLLLTLDLSN-TEYELGKMSAIALEMIHTYSLIHDPLA 86
DB 25 SSRESVLSYIHAGGKRIRPFLLEVLLEALQVTKPAHAQVATALEMIHTGSLIHDPLA 84
QY 87 MDNDYRRGKLTNHNKVGWETAILAGDALITKAFELISSDDRLTDEVKIKVLRSLASG 146
DB 85 MDDDDYRRGLTNHNKFGEMAILAGDALFDLPYALIAQAD-LPSQIKVDLIANLSLASG 143
QY 147 HVGWVGQMLDMQSEGPIDLETLEMIHKTGTGALLTFVAVMSAADIANVDDTTKEHLESY 206
DB 144 SLGWAGQVLDMEGEHQHLSLEELQITTHANKTKLLAYPFOAAAIIAELSPENQVRLKT 203
QY 207 SYHLGMMFOIKDLDLCYGDGDEAKLGKVGSDLENNKSTYVSLGKGDGAEDKLTTHRDAV 266
DB 204 GELIGLAFVRDDVLDVTASFEGIKTPQKDLQAEKSTYVLPALLGLEESIAFCNQTLDOAN 263
267 DELTQIDEQ--FNTKHLLEIVD 286
DB 264 EKLEETAAQLPFTTESIVSVE 285

RESULT 7

US-09-815-242-13273
; Sequence 13273, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13273
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13273

Query Match 37.6%; Score 547; DB 10; Length 291;
Best Local Similarity 45.8%; Pred. No. 3.4e-40;
Matches 120; Conservative 50; Mismatches 88; Indels 4; Gaps 3;

QY 28 TQLEESMLYSLNAGGKRIRPVLLLTLDLSN-TEYELGKMSAIALEMIHTYSLIHDPLA 86
DB 25 SSRESVLSYIHAGGKRIRPFLLEVLLEALQVTKPAHAQVATALEMIHTGSLIHDPLA 84

QY 87 MDNDYRRGKLTNHNKVGWETAILAGDALITKAFELISSDDRLTDEVKIKVLRSLASG 146
DB 85 MDDDDYRRGLTNHNKFGEMAILAGDALFDLPYALIAQAD-LPSQIKVDLIANLSLASG 143
QY 147 HVGWVGQMLDMQSEGPIDLETLEMIHKTGTGALLTFVAVMSAADIANVDDTTKEHLESY 206
DB 144 SLGWAGQVLDMEGEHQHLSLEELQITTHANKTKLLAYPFOAAAIIAELSPENQVRLKT 203
QY 207 SYHLGMMFOIKDLDLCYGDGDEAKLGKVGSDLENNKSTYVSLGKGDGAEDKLTTHRDAV 266
DB 204 GELIGLAFVRDDVLDVTASFEGIKTPQKDLQAEKSTYVLPALLGLEESIAFCNQTLDOAN 263
267 DELTQIDEQ--FNTKHLLEIVD 286
DB 264 EKLEETAAQLPFTTESIVSVE 285

RESULT 8

US-09-941-947A-20
; Sequence 20, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzostowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odom, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL1903 US NA
; CURRENT APPLICATION NUMBER: US/09/941,947A
; PRIOR FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Methylobionas 16a
US-09-941-947A-20

Query Match 35.5%; Score 516.5; DB 9; Length 297;
Best Local Similarity 45.0%; Pred. No. 1.6e-37;
Matches 113; Conservative 47; Mismatches 88; Indels 3; Gaps 2;

QY 23 KSVMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSN-TEYELGKMSAIALEMIHTYSLIHD 82
DB 26 ENILPOTLQAMRYSVLNGKRTPLTYATGALGPENVLDPACAVEFIHVYSLIHD 85
QY 83 DLPAMDNDYRRGKLTNHNKVGWETAILAGDALITKAFELISSDDRLTDE--VKIKVLR 140
DB 86 DLPAMDNDLRRGKPTCHKAYDEATAILAGDALQALAFEVLANDPGITVDAPARKMITA 145
QY 141 LSTASGHVGVGQMLDMQSEGPIDLETLEMIHKTGTGALLTFVAV-MSAADIANVDDTT 199
DB 146 LTRASGQGVGQADLDGSGRKLTLPELNNHHTKALIRASVLAALSXPDLDTCV 205
QY 200 KEHLESYSHLGLMMFOIKDLDLCYGDGDEAKLGKVGSDLENNKSTYVSLGKGDGAEDKLT 259
DB 206 AKKLDHYAKCIGLSFQVKDDILDIEDATATLGTQGTQKDDINDKPTTPALLGMAKQAKQ 265
QY 260 YHRDAVDEL 270
DB 266 ELHEQAVESIT 276

RESULT 9

QY 13 VNNELSVAINKSVMT-QLEESMLYSLNAGGRIRPVLVLLTLDLSLNTYEYELGMSAIAL 71
Db 89 VNOALDAAI--ALRDPHKIHOAMRYSLLAGGRVRPVLICIRACELVGGTEATAIPAACAV 146
QY 72 EMHTYSLIHDLPAMDNDYRRGKLTNKHVYGEWTAILAGDALLTKAFE-LISSDDRIT 130
Db 147 EMHTYSLIHDLPAMDNDLARGPTNKHVYGEVDVAVLAGDALLAFEFHVAATEGVS 206
QY 131 DEVKIKVQLRSLASGHVGMGOMLDMQSEG-QPDLTLEMIHKTGTGALLTFAYMSA 189
Db 207 PSRVRAIGELAKSIGTEGLVAGQVVDIDSEGVANVLETLEFIHVHKTAALEAAVVLG 266
QY 190 ADIANVDDTTREHLESYSYHLGMMFOIKDDLLDCYDEAKLGKKGVSLENKNSYVSL 249
Db 267 AIVGGSDVEEVEKLRKARFCIGLLFQVVDIILDVTKSSSELGKTAGKDLVADKVTPKLL 326
QY 250 GKDAED 256
Db 327 GIDKSKE 333

RESULT 15

US-09-815-242-10069
; Sequence 10069, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10069

; LENGTH: 299

; TYPE: PRF

; ORGANISM: Escherichia coli

US-09-815-242-10069

Query Match 33.4%; Score 485.5; DB 10; Length 299;
Best Local Similarity 41.8%; Pred. No. 8.6e-35;
Matches 118; Conservative 42; Mismatches 117; Indels 5; Gaps 4;

QY 10 IDEVNNELSVAINK-SVMDTQLEESMLYSLNAGGRIRPVLVLLTLDLSLNTYEYELGMSA 68
Db 11 VQANOALSRFTAPLPFQNTPVETMQYCALIGGKRLRPFLVYATGFMFGVSTNTLDAPA 70
QY 69 TALEMIHYSLIHDLPAMDNDYRRGKLTNKHVYGEWTAILAGDALLTKAFELISSDD- 127
Db 71 RAVECIHAYSLIHDLPAMDNDLRRGLPTCHVKEGEANAILAGDLQTLAFSILSDADM 130

QY 128 -RLTDEVKIKVQLRSLASGHVGMGOMLDMQSEGQPIDLETLEMIHKTGTGALLTFAY 186
Db 131 PEVSDRDRISMISLASASGIACGGQALDLDABEGKHVPLDALERIHRHKTGALTRAAY 190
QY 187 -MSAADIANVDDTTREHLESYSYHLGMMFOIKDDLLDCYDEAKLGKKGVSLENKNSY 245
Db 191 RLGALSAGDKGRRALPVLVDKYAESIGLAFQVODDILDVVGDVATLGKROGADOQLGKSY 250
QY 246 VSLGKDGAEKLTVHRDAAVDELTDQIDQ-FNTKHLLEIVD 286
Db 251 PALLGLEQARKKARDLIIDARQSLKQAEQSLDTSALEALAD 292

Search completed: May 29, 2003, 12:08:17
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: May 29, 2003, 11:56:10 ; Search time 19 Seconds
(without alignments)
1452.136 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPKMLIDEVNNELSA.....ELTQIDQFNKHLLEIVDL 287
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1444	99.4	293	2 A89932	hypothetical prote
2	702.5	48.3	293	2 AC1245	geranyltransferase
3	686.5	47.2	293	2 AG1607	geranyltransferase
4	654	45.0	297	2 JX0257	geranyltransferase
5	638	43.9	294	2 E83997	geranyltransferase
6	636.5	43.8	272	2 A89961	geranyltransferase
7	549	37.8	291	2 G95139	geranyltransferase
8	547	37.6	291	2 F98007	geranyltransferase
9	540.5	37.2	285	2 A86732	geranyltransferase
10	534.5	36.8	309	2 AE1833	geranyltransferase
11	532	36.6	289	2 D97156	probable geranylge
12	530.5	36.5	300	2 T06969	farnesyltransferase
13	526.5	36.2	302	2 S74538	geranylgeranyl pyr
14	497.5	34.2	295	2 C64123	geranyltransferase
15	489.5	33.7	357	2 T09966	farnesyltransferase
16	485.5	33.4	298	2 E81796	geranyltransferase
17	485.5	33.4	299	2 J00665	geranyltransferase
18	485.5	33.4	299	2 C30688	geranyltransferase
19	485.5	33.4	299	2 G85538	geranyltransferase
20	484	33.3	294	2 A82267	geranyltransferase
21	479.5	33.0	369	2 S53722	farnesyltransferase
22	479	33.0	366	2 T10452	farnesyltransferase
23	478.5	32.9	316	2 T11021	farnesyltransferase
24	478	32.9	371	2 F85434	geranylgeranyl pyr
25	475.5	32.7	306	2 AH0385	geranyltransferase
26	474.5	32.7	295	2 F83139	geranyltransferase
27	473	32.6	299	2 AG0554	geranyltransferase
28	470.5	32.4	294	2 D87505	geranyltransferase
29	469.5	32.3	259	2 F81217	geranyltransferase

30	467	32.1	347	2 E84566	probable geranylge
31	457.5	31.5	304	2 F97685	geranyltransferase
32	457.5	31.5	335	2 AG2910	geranyltransferase
33	438	30.1	327	2 D64407	bifunctional short
34	436	30.0	372	2 G84566	probable geranylge
35	434.5	29.9	304	2 A13285	geranyltransferase
36	434	29.9	376	2 T02429	geranylgeranyl pyr
37	419	28.8	378	2 S71230	geranylgeranyl pyr
38	417	28.7	281	2 B81261	geranyltransferase
39	405	27.9	291	2 D82778	geranyltransferase
40	378	26.0	325	2 T45152	bifunctional short
41	376.5	25.9	302	2 A37802	crIE protein - Ew
42	375	25.8	282	2 B84984	geranyltransferase
43	373.5	25.7	303	2 A64636	geranyltransferase
44	372.5	25.6	303	2 G71878	geranyltransferase
45	372	25.6	346	2 C75139	bifunctional short

ALIGNMENTS

RESULT 1

A89932
hypothetical protein ispA [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: A89932
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89932
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <KUR>
A:Cross-references: GB:BA000018; PID:g13701320; PIDN:BAB42614.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ispA

Query Match	99.4%	Score 1444	DB 2	Length 293
Best Local Similarity	99.3%	Pred. No. 1.9e-99		
Matches 285	Conservative 1	Mismatches 1	Indels 0	Gaps 0
QY	1	MTNLPKMLIDEVNNELSAVINKSVMDTOLESMLYSLNAGGKRIRPVLILLTLDLSLNT	60	
DB	1	MTNLPKMLIDEVNNELSAVINKSVMDTOLESMLYSLNAGGKRIRPVLILLTLDLSLNT	60	
QY	61	YELGMKSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHHKYGEWTAIAGDALITKAF	120	
DB	61	YELGVKSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHHKYGEWTAIAGDALITKAF	120	
QY	121	ELISSDDRLTDEVKIKVLRSLASGHVGVGQMLDMQSEGQPDLETFLEMIHKTGA	180	
DB	121	ELISSDDRLTDEVKIKVLRSLASGHVGVGQMLDMQSEGQPDLETFLEMIHKTGA	180	
QY	181	LITFAVMSAADIANVDDTTKEHLESYSHLGAMFQIKDLDLCYGEAKLKKVGSGLN	240	
DB	181	LITFAVMSAADIANVDDTTKEHLESYSHLGAMFQIKDLDLCYGEAKLKKVGSGLN	240	
QY	241	NKSTVYSLGKGAEDKLTYYHRDAAVDELDTQIDEQFNTKHLLEIVDL	287	
DB	241	NKSTVYSLGKGAEDKLTYYHRDAAVDELDTQIDEQFNTKHLLEIVDL	287	

RESULT 2

AC1245
geranyltransferase homolog lmo1363 [imported] - Listeria monocytogenes (strain E
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C:Accession: AC1245

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1245
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99441.1; PID:g16410779; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lml363
C:Superfamily: geranyltransferase

Query Match 48.3%; Score 702.5; DB 2; Length 293;
Best Local Similarity 52.9%; Pred. No. 1.2e-44;
Matches 148; Conservative 46; Mismatches 81; Indels 5; Gaps 2;

QY 8 KLIDEVNNELSVKSVMDTOLESMLYSLNAGGKRIRPVLLLLDLSLNTYEYELGMS 67
Db 13 KVDE---SLFKEINMRNIEPKLESMLYSVQAGGKRIRPMLVFATLQALNPIEMRGLKT 69
QY 68 AIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALLTKAFELISSDD 127
Db 70 AIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALLTKAFELISSDD 129
QY 128 RLDEVKIKVQLRSLSTASGHVGMVGMQMDSEGQPIDLETLMHTKTKGALLTFVYM 187
Db 130 NLSFETRIALINQISYSSGAEGMVGGLADLEAKENQVTEELSSIHARKTGELLIVAVT 189
QY 188 SAADIANVDDTTKEHLESYSHGLMMFOIKDDLLDCYDEAKLGKGVSDLENNKSTYVS 247
Db 190 SAAKIAEAPQPKRLRIFAENIGIGFQISDDILDVIGDETKMGKKTGDAFLNKSTYGP 249
QY 248 LLGKGAEDKLTVHRDAAVDELQIDQFNTHLLEIVDL 287
Db 250 LLTLDGAKRALNEHVHTIAKSALSCHD--FDDEILLKLADL 287

RESULT 3
AG1607
geranyltransferase homolog lin1400 [Imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
Accession: AG1607
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1607
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96631.1; PID:g16413873; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin1400
C:Superfamily: geranyltransferase

Query Match 47.2%; Score 686.5; DB 2; Length 293;
Best Local Similarity 52.1%; Pred. No. 1.8e-43;
Matches 146; Conservative 45; Mismatches 84; Indels 5; Gaps 2;

QY 8 KLIDEVNNELSVKSVMDTOLESMLYSLNAGGKRIRPVLLLLDLSLNTYEYELGMS 67
Db 13 KVDE---SLFKEINMRNIEPKLESMLYSVQAGGKRIRPMLVFATLQALNPIEMRGLKT 69

Db 13 KVDE---SLFKEINMRNIEPKLESMLYSVQAGGKRIRPMLVFATLQALNPIEMRGLKT 69
QY 68 AIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALLTKAFELISSDD 127
Db 70 AIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALLTKAFELISSDD 129
QY 128 RLDEVKIKVQLRSLSTASGHVGMVGMQMDSEGQPIDLETLMHTKTKGALLTFVYM 187
Db 130 NLSFETRIALINQISYSSGAEGMVGGLADLEAKENQVTEELSSIHARKTGELLIVAVT 189
QY 188 SAADIANVDDTTKEHLESYSHGLMMFOIKDDLLDCYDEAKLGKGVSDLENNKSTYVS 247
Db 190 SAAKIAEAPQPKRLRIFAENIGIGFQISDDILDVIGDETKMGKKTGDAFLNKSTYGP 249
QY 248 LLGKGAEDKLTVHRDAAVDELQIDQFNTHLLEIVDL 287
Db 250 LLTLDGAKRALNEHVHTIAKSALSCHD--FDDEILLKLADL 287

RESULT 4
JX0257
geranyltransferase (EC 2.5.1.10) - *Bacillus stearothermophilus*
N:Alternate names: farnesyl-diphosphate synthase
C:Species: *Bacillus stearothermophilus*
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
C:Accession: JX0257
R;Koyama, T.; Obata, S.; Osabe, M.; Takeshita, A.; Yokoyama, K.; Uchida, M.; Nishino,
J.; Biochem. 113, 355-363, 1993
A:Title: Thermotable farnesyl diphosphate synthase of *Bacillus stearothermophilus*: m
A:Reference number: JX0257; MUID:93252758; PMID:8486607
A:Accession: JX0257
A:Molecule type: DNA
A:Residues: 1-297 <KOY>
A:Cross-references: GB:D13293; NID:g391609; PIDN:BAA02551.1; PID:g391610
C:Genetics:
A:Start codon: GTG
C:Superfamily: geranyltransferase
C:Keywords: transferase

Query Match 45.0%; Score 654; DB 2; Length 297;
Best Local Similarity 47.7%; Pred. No. 4.8e-41;
Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;

QY 1 MTNLPNKLIDEVNNELSVKSVMDTOLESMLYSLNAGGKRIRPVLLLLDLS 56
Db 1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLLKAMAYSLGAGGKRIRPRLLSVRA 60
QY 57 LNTYEYELGMSKSALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALL 116
Db 61 LGKDPVGLPVACIAEMHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALL 120
QY 117 TRAFELIS--SDRLTDEVKIKVQLRSLSTASGHVGMVGMQMDSEGQPIDLETLMHT 174
Db 121 TYAFQLITEDDERIPPSVRLRIERLAKAAGPEGVAGQADMEGKTLTLESEYTH 180
QY 175 KTKTGALLTFVYMSAADIANVDDTTKEHLESYSHGLMMFOIKDDLLDCYDEAKLGKV 234
Db 181 RHKTGKMLQYSHVAGALIGGADARQTRLEDEFAHLGLAFQIRDDILDIEGAEKIGKV 240
QY 235 GSDLENNKSTYVSLLGKGAEDKLTVHRDAAVDELQID 273
Db 241 GSDQSNKATYPALLSLAGAKELAPHIEAARHLRNAD 279

RESULT 5
E83997
geranyltransferase BH2781 [Imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 17-May-2002
C:Accession: E83997
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: E83997

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-294 <STO>

A:Cross-references: GB:AF001516; GB:BA000004; NID:g10175192; PIDN:BA066500.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2781

C:Superfamily: geranyltransferase

Query Match 43.9%; Score 638; DB 2; Length 294;

Best Local Similarity 47.7%; Pred. No. 7,1e-40;

Matches 134; Conservative 53; Mismatches 92; Indels 2; Gaps 1;

QY 6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLLLLFLDLSNTEYELGM 65

DB 9 LDEIKDIIEERMPAHERLNSPDLKNSMLYSLRAGGRIRPALLATLMKSFQKDISQGI 68

QY 66 KSAIALEMIHYSLIHDDLPAMDNDYRGRKLTNKHVYGEWTAIAGDALLTKAFELISS 125

DB 69 DLACAIEMIHYSLIHDDLPMSDDDIRRGKPTNKHVGEAHAILAGDALLTYSFEIVAK 128

QY 126 DDLRTDEVKIVLQRLSTASGHVGMVGMQMDQSEGQPIDLETLEMIHKTGTGALLTFA 185

DB 129 MKGVDPATLGLIELARAGPEGVGMVGMQVADIEGKNKLVGELEYHHHKTGALLSFA 188

QY 186 VMSAADIANDVDTTKEHLESYHYHGMFQIKDDLLDCYGEAKLGRKVGSDLENKSTY 245

DB 189 IVAGARLADASEQDIENIRRESRELGLFQIKDDLLDVEGDAQAIGKPVSDGNGKSTY 248

QY 246 VSLIGKCAEAKLTHYRDAANDVDELTDQIDQENKHLLEIVD 286

DB 249 PSLLTLEGAKELHLTLAKKEYLSV--QMNHLLEELTD 287

RESULT 6

A89961

geranyltransferase homolog yqjD - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: A69961

A:Authors: F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillette, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall

ier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Author: P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols

A.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanton

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: A69961

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-272 <KUN>

A:Cross-references: GB:Z99116; GB:AL009126; NID:g2634723; PIDN:CAB14359.1; PID:g2634862

A:Experimental source: strain 168

C:Genetics:

A:Gene: yqjD

C:Superfamily: geranyltransferase

Query Match 43.8%; Score 636.5; DB 2; Length 272;

Best Local Similarity 49.3%; Pred. No. 8.3e-40;

Matches 134; Conservative 51; Mismatches 82; Indels 5; Gaps 2;

QY 1 MTNLPNNKLIDE---VNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLLLLTDSL 57

DB 1 MTNKLTSFLADRKKTENQLSVYTEKLDMPDLSKKSMYLSQAGGRRLPLVLAVLNAY 60

QY 58 NTEYELGKMSAIALEMIHYSLIHDDLPAMDNDYRGRKLTNKHVYGEWTAIAGDALIT 117

DB 61 GKSEKDGIPVGCACAVEMIHYSILIHDDLPAMDNDYRGRKLTNKHVYGEWTAIAGDGLIT 120

QY 118 KAFELISS--DDRLTDEVKIVLQRLSTASGHVGMVGMQMDQSEGQPIDLETLEMIHK 175

DB 121 ESFKLITSHVSDVSAEKRLRLVNLISAAGTEGMVGMQVADMEAGNRQVTFELESIEH 180

QY 176 TKTGALLTFAYMSAADIANDVDTTKEHLESYHYHGMFQIKDDLLDCYGEAKLGRKVG 235

DB 181 KTSKLLGFCVAGAILADAPEDTETLTFSSHIGIGFQIRDDIIDLLEGSEKIGKRVG 240

QY 236 SLENKSTYVYSLIGKDGAEKLTLYHRAAVD 267

DB 241 SDTTNDKSTYPSLLSLEGAKHKLDVHKEAND 272

RESULT 7

G95139

geranyltransferase [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C:Accession: G95139

R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H

on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl

son, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris

A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: G95139

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-291 <KUR>

A:Cross-references: GB:AE005672; PIDN:AK75312.1; PID:g14972686; GSPDB:GN00164; TIGR:

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP1205

Query Match 37.8%; Score 549; DB 2; Length 291;

Best Local Similarity 46.2%; Pred. No. 2.7e-33;

Matches 121; Conservative 50; Mismatches 87; Indels 4; Gaps 3;

QY 28 TOLESMLYSLNAGGRIRPVLLLLTDSL--TEVELGKMSAIALEMIHYSLIHDDLP 86

DB 25 SSLRESVLSYIHAGGRIRPFLLLEVLQVTKIRPAHQVATALEMIHYSLIHDDLP 84

QY 87 MDNDYRGRKLTNKHVYGEWTAIAGDALLTKAFELISSDDRLTDEVKIVLQRLSIASG 146

DB 85 MDDDDYRGRLTNKHVYGEWTAIAGDALLTKAFELISSDDRLTDEVKIVLQRLSIASG 143

QY 147 HVGVMVGMQMDQSEGQPIDLETLEMIHKTGTGALLTFAYMSAADIANDVDTTKEHLESY 206

DB 144 SLUGVAGQVLDMEGEHQLSLSELOTHANKTKGLLAYPFAAAITAEUSPEMKVTKTV 203

QY 207 SYHLGMFQIKDDLLDCYGEAKLGRKVGSDLENKSTYVYSLIGKDGAEKLTLYHRAAV 266

DB 204 GELIGLAFQVRDVLDTVTSFEEIGTKPQKDLQAEKSTYPALLGLEESIAFCNQITLDQAN 263

QY 267 DELTQIDQ--FWTKHLEIVD 286

DB 264 EKLEETAQQLPPESTESIVSVE 285

RESULT 8

F98007

geranyltransferase (EC 2.5.1.10) [imported] - Streptococcus pneumoniae (strain R

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001

C:Accession: F98007

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <RUP>
A:Cross-references: GB:AE001437; PIDN:AAK80039.1; PID:g15025068; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2080

Query Match 36.6%; Score 532; DB 2; Length 289;

Best Local Similarity 44.9%; Pred. No. 4.8e-32;

Matches 127; Conservative 44; Mismatches 108; Indels 4; Gaps 3;

QY 7 NKLI-DEVNELSVAI-NKSVMDTOLESMLYSLNAGGRKRIRPVLLLTLDLSLNEYLG 64

DB 3 NKVIKEVEEYLSRYPEGKDNKRVYEMSYSLNAGGRKRIRPVLLLTLDLSLNEYLG 62

QY 65 MKSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALITKAFELIS 124

DB 63 IDIAAAIEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALITKAFELIS 122

125 SDDRLTDEVKIKVLRSLASGHVGMQGLDMQSEGQIDLETLEMIHKTGALLTF 184

123 KYICGKGAEALACIMISKAASSDGMIGGVVDILSEGGKINEDRLYMHKKKTGELIKA 182

QY 185 AVMSAADIANVDDTTKEHLESYSYHLGMMFOIKDDLLDCYGEAKLGKVGSDLENNKST 244

DB 183 AVVSAIIGGAPLHDVVELLSQYDKLGLAFQEDDILDIGTKMGTKSDLENDKCT 242

QY 245 YVSLGKDGAEDEKLTTHYRDAAVDELTDQIDQFNKHLLEIVDL 287

DB 243 YVTLGIDKCKSICRELTDCLDIIGKI--QGTELLKEITEL 283

RESULT 12

T06969

farnesyltransferase (EC 2.5.1.29) - Cyanophora paradoxa cyanelle

C:Species: cyanelle Cyanophora paradoxa

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 24-Sep-1999

C:Accession: T06969

R:Stirewait, V.I.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.

submitted to the EMBL data library, July 1995

A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.

A:Reference number: 215840

A:Accession: T06969

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-300 <ST>

A:Cross-references: EMBL:U30821; NID:g1016083; PIDN:AAA81312.1; PID:g1016225

Experimental source: strain Pringsheim LB555

Genetics:

Gene: crtE

A:Genome: cyanelle

C:Superfamily: geranyltransferase

C:Keywords: carotenoid biosynthesis; chromoplast; cyanelle; transferase

Query Match 36.5%; Score 530.5; DB 2; Length 300;

Best Local Similarity 43.0%; Pred. No. 6.5e-32;

Matches 120; Conservative 51; Mismatches 103; Indels 5; Gaps 3;

QY 2 TNLPMKLIDEVNELSVAINKSV---MDTOLESMLYSLNAGGRKRIRPVLLLTLDLSLN 58

DB 6 TTNELTYLIERKEIVEDTLNKSIPRGNPTFYDSIRYSLSAGGRKRIRPILCLASCEL 65

QY 59 TVEYELGMSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALITK 118

DB 66 GTMENALPACALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALIT 125

QY 119 APELSSDDR-LTDEVKIKVLRSLASGHVGMQGLDMQSEGQ-PIDLETLEMIHKT 176

DB 126 APEFIATQKNYPADLVIVKVAHSVTTSLGVGGQIIDLSSECKSDTTLTFNFIH 185

QY 177 KTGALLTFVMSAADIANVDDTTKEHLESYSYHLGMMFOIKDDLLDCYGEAKLGKVG 236

Db

QY

Db

RESULT 13

S74538

geranylgeranyl pyrophosphate synthase - Synechocystis sp. (strain PCC 6803)

N:Alternate names: hypothetical protein slr-0739

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S74538

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys

S.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74538

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-302 <KAN>

A:Cross-references: EMBL:D90899; GB:AB001339; NID:g1515650; PIDN:BAAL6690.1; PID:g165

A:Note: The nucleotide sequence was submitted to the EMBL data Library, June 1996

C:Superfamily: geranyltransferase

Query Match 36.2%; Score 526.5; DB 2; Length 302;

Best Local Similarity 44.2%; Pred. No. 1.3e-31;

Matches 117; Conservative 49; Mismatches 94; Indels 5; Gaps 3;

QY 4 LPMNKLIDEVNELSVAINKSVMDTOLESMLYSLNAGGRKRIRPVLLLTLDLSLNEYEL 63

DB 16 LQVKKGVVEAALDSSLAIR---PEKIYEMRYSLLAGGRKRIRPILCIACELCGGDEAL 72

QY 64 GKMSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALITKAFELI 123

DB 73 ALPTACALEMIHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAIAGDALITKAFELI 132

QY 124 SDDRLTDEVKIKVLRSLASGHVGMQGLDMQSEGQ-PIDLETLEMIHKTGALLTF 181

DB 133 VHTPTQADPQALLQVIRLGRVVGAGLVGGQVLDSEGRDITPETFTHHTKTGAL 192

QY 182 LTFVMSAADIANVDDTTKEHLESYSYHLGMMFOIKDDLLDCYGEAKLGKVGSDLENN 241

DB 193 LEASVLTGAILAGATGEQOQLARYAQNIGLAFQVVDLITATQEEELGKTAGDKVKAQ 252

QY 242 KSTYVSLGKDGAEDEKLTTHYRDAAV 266

DB 253 KATYPSLLGLEASRAQASLIDQAI 277

RESULT 14

C64123

geranyltransferase (EC 2.5.1.10) - Haemophilus influenzae (strain Rd KW20)

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999

C:Accession: C64123

R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: C64123

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-295 <TIG>

A:Cross-references: GB:U32822; GB:L42023; NID:g1574265; PIDN:AAC23087.1; PID:g1574277

C:Genetics:

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:52:16 ; Search time 11 Seconds
(without alignments)
1082.155 Million cell updates/sec

Title: US-09-925-637-64

Perfect score: 1453

Sequence: 1 MTNLPKMLIDEVNNELSA.....ELTQIDEQFNKHLLEIVDL 287

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	654	45.0	297	1	ISPA_BACST
2	649	44.7	291	1	ISPA_MICLU
3	636.5	43.8	272	1	ISPA_BACSU
4	530.5	36.5	300	1	CRPE_CYPAP
5	497.5	34.2	295	1	ISPA_HAEIN
6	489.5	33.7	357	1	GGPP_CATRO
7	485.5	33.4	299	1	ISPA_ECOLI
8	479.5	33.0	369	1	GGPP_CAPAN
9	479	33.0	366	1	GGPP_SINAL
10	478	32.9	371	1	GGPP_ARATH
11	438	30.1	327	1	ISPA_METJA
12	376.5	25.9	302	1	CRPE_PANAN
13	376.5	25.9	324	1	ISPA_METTM
14	375	25.8	282	1	ISPA_BUCAI
15	352	24.2	332	1	ISPA_RHISN
16	346	23.8	307	1	CRPE_ERWHE
17	345.5	23.8	325	1	ISPA_WETH
18	338	23.3	332	1	GGPP_SULSO
19	334	23.0	332	1	ISPA_BRAJA
20	330.5	22.7	362	1	ISPA_AQUAE
21	330	22.7	330	1	GGPP_SULAC
22	320.5	22.1	289	1	CRPE_RHOCA
23	311	21.4	323	1	ISPB_ECOLI
24	302.5	20.8	288	1	CRPE_RHOSH
25	294	20.2	329	1	ISPB_HAEIN
26	273	18.8	359	1	GGPP_MYCTU
27	260.5	17.9	348	1	HEP2_BACSU
28	254	17.5	323	1	PREA_PORPU
29	250	17.2	320	1	HEP2_BACST
30	248.5	17.1	323	1	PREA_CYPAP
31	245.5	16.9	323	1	PREA_SYNY3
32	211	14.5	323	1	PREA_CYACA
33	201.5	13.9	378	1	DPS_SCHPO

34	197.5	13.6	473	1	COOL_YEAST
35	173.5	11.9	352	1	FPPS_YEAST
36	158.5	10.9	347	1	FPPS_SCHPO
37	156.5	10.8	367	1	FPPS_CHICK
38	149.5	10.3	353	1	FPPS_RAT
39	149	10.3	349	1	FPPS_KLULA
40	147.5	10.2	300	1	GGPP_HUMAN
41	147.5	10.2	343	1	FPPS_ARTAN
42	141.5	9.7	294	1	GGPP_BOVIN
43	141.5	9.7	384	1	FPPL_ARATH
44	141	9.7	300	1	GGPP_MOUSE
45	141	9.7	353	1	FPPS_HUMAN

ALIGNMENTS

RESULT 1
ISPA_BACST STANDARD; PRT; 297 AA.
ID Q08291; Q53435; Q53436; Q53437; Q53438;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Geranyltransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE (fpp synthase).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-ATCC 10149;
RX MEDLINE=93252758; PubMed=8486607;
RA Koyama T., Obata S., Osabe M., Takeshita A., Yokoyama K.,
R Uchida M., Nishino T., Ogura K.;
RT "Thermostable farnesyl diphosphate synthase of Bacillus
RT stearothermophilus: molecular cloning, sequence determination,
RT overproduction, and purification.";
RL J. Biochem. 113:355-363(1993).
RN [2]
RP MUTAGENESIS OF CYSTEINE RESIDUES.
RX MEDLINE=95001990; PubMed=7918490;
RA Koyama T., Obata S., Saito K., Takeshita-Koike A., Ogura K.;
RT "Structural and functional roles of the cysteine residues of Bacillus
RT stearothermophilus farnesyl diphosphate synthase.";
RL Biochemistry 33:12644-12648(1994).
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC = diphosphate + trans,trans-farnesyl diphosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic
CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to licensed@isb-sib.ch).
CC
CC EMBL: D13293; BAA02551.1;
DR EMBL: S72629; AAB32272.1;
DR EMBL: S72630; AAB32273.2; ALT_SEQ.
DR EMBL: S72633; AAB32274.1;
DR EMBL: S72635; AAB32275.2; ALT_SEQ.
DR PIR: JX0257; JX0257.
DR InterPro: IPR000092; Polyprenyl_synt.
DR Pfam: PF00348; polyprenyl_synt_1;
DR PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
DR PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
KW Transferase; Isoprene biosynthesis.
FT MUTAGEN 289 73 C->F,S: NO LOSS OF ACTIVITY.
FT MUTAGEN 289 73 C->F,S: NO LOSS OF ACTIVITY.
SQ SEQUENCE 297 AA; 32310 MW; 0F921C3F029EBB6 CRC64;

Query Match 45.08; Score 654; DB 1; Length 297;
Best Local Similarity 47.7%; Pred. No. 3.9e-41;
Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;

QY 1 MTNPLMKLDEVNELSVAINKSVMD-----DTQLESMYSLNAGGRIRPVLLTLDLSTNTEYELGK 66
DB 1 MAQSVQFLNEQQAQVETALSRYIERLEGPKALKKAMAYSLEAGGRIRPVLLTLDLSTVRA 60
QY 57 LNTVEYELGKMSAJALEMHIYSLTHDDLPAWMDNDYRGRKLTNNHKVYGEWTAIAGDALL 116
DB 61 LGRDPVGLPVACAIEMHIYSLTHDDLPAWMDNDLGRPTNNHKVYGEWTAIAGDALL 120
QY 117 TKAFELIS--SDDRLTDEVKIKVQLRSLASGHVGMVGGQMLDMSQEGQPIDLETLEMHKTKYTGALLTF 174
DB 121 TYAFOLITEIDDERIPPVRLRLIERLAKAAGPBGVAGQAADMEGEGTKLTLSLEYIH 180
QY 175 KTKTGALLTFVMSAADIANDVDTTKEHLESYSHLGMMFOIKDDLDYCGDEAKKGVSGDLENKSTYV 234
DB 181 RHKTGKMLQSVHAGALIGGADARQTRDELDEFAHLGLAFQIRDDIILDIEGAEEKIGPV 240
QY 235 GSDLENKSTYVSLGKDGAEKLTYYHRDAVAVDLTQID 273
DB 241 GSDSNKNTYPPALLSLAGAKELAFHIEAAQRHLRNAD 279

RESULT 2
ID ISPA_MICLU STANDARD; PRT; 291 AA.
AC O66126; Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Geranyltransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE (PPP synthase).
GN EPS.
OS Micrococcus luteus (Micrococcus lysodeikticus).
OC Bacteria; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Micrococcus.
OX NCBI_TaxID=1270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B-P 26;
RX MEDLINE=98175686; PubMed=9515931;
RA Shimizu N., Koyama T., Ogura K.;
RT "Molecular cloning, expression, and characterization of the genes
encoding the two essential protein components of Micrococcus luteus
B-P 26 hexaprenyl diphosphate synthase."
J. Bacteriol. 180:1578-1581(1998).
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
= diphosphate + trans,trans-farnesyl diphosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AB003187; BAA25265.1;
CC InterPro; IPR000092; Polyrenyl_synth.
CC Pfam; PF00348; polyrenyl_synth. 1.
CC PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
CC PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
CC TRANSFERASE; Isoprene biosynthesis.
CC SQ SEQUENCE 291 AA; 32362 MW; 797125AB71E5674A CRC64;

Query Match 44.7%; Score 649; DB 1; Length 291;
Best Local Similarity 52.9%; Pred. No. 8.9e-41;
Matches 146; Conservative 40; Mismatches 78; Indels 12; Gaps 4;

QY 13 VNNELSVAINKSVMD-----TQLESMYSLNAGGRIRPVLLTLDLSTNTEYELGK 66
DB 8 MNRDFNLINELSLNKKYHPAQSRHLHEAINYSLSAGGRIRPVLLTLDLSTNTEYELGK 67
QY 67 SAIALEMHIYSLTHDDLPAWMDNDYRGRKLTNNHKVYGEWTAIAGDALLTKAFELISSD 126
DB 68 FGIALEMHIYSLTHDDLPAWMDNDYRGRKLTNNHKVYGEWTAIAGDALLTKAFELISSD 126
QY 127 DRLTDEVKIKVQLRSLASGHVGMVGGQMLDMSQEGQPIDLETLEMHKTKYTGALLTF 186
DB 127 TQLENAETKLSLNLSTASGNGMYGQMLDMSQEGQPIDLETLEMHKTKYTGALLTF 186
QY 187 MSAADIANDVDTTKEHLESYSHLGMMFOIKDDLDYCGDEAKKGVSGDLENKSTYV 246
DB 187 VSAIINFNDAQIEQLNIIGKNVGLMFQIKDDLDVGEFENIGKTVGSDLNNDKSTYV 246
QY 247 SLGKDGAE-----DKLTYHRDAVAVDLTQIDQFNT 278
DB 247 SLGKDGAEKLTYYHRDAVAVDLTQIDQFNT 281

RESULT 3
ID ISPA_BACSU STANDARD; PRT; 272 AA.
AC P54383;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Geranyltransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE (PPP synthase).
GN YQID.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
the Bacillus subtilis genome containing the skin element and many
sporulation genes."
Microbiology 142:3103-3111(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Deniot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Colightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Huilo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OX Haemophilus.
 RN NCBI_TaxID=727;
 RP [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-Rd / KW20 / ATCC 51907;
 RA MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
 CC = diphosphate + trans-trans-farnesyl diphosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U32822; AAC23087.1;
 DR TIGR; H11438;
 DR InterPro; IPR000092; Polyrenyl_synth.
 DR Pfam; PF00348; polyrenyl_synth; 1.
 DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
 DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
 KW transferase; isoprene biosynthesis; Complete proteome.
 SQ SEQUENCE 295 AA; 32407 MW; A64DE650266E670F CRC64;
 Query Match 34.2%; Score 497.5; DB 1; Length 295;
 Best Local Similarity 41.0%; Pred. No. 1.2e-29;
 Matches 116; Conservative 52; Mismatches 100; Indels 15; Gaps 4;
 QY 16 ELSVAINKSVMDQLE-----ESMLYSLNAGGKRIRPVLLTLDLSLNTVEYELGM 65
 10 QVQTRINR-FLEAQFEGIESHNAPLEANKYALLGGKVRPFLVYATQGLCAEKQTLTD 68
 66 KSAIALEMTHTVSLIHDDLPAMDNDYRRGKLTNHRKVGWETAILAGDALLKAFELISS 125
 69 YAAAEIAEIAHAYSLIHDDLPAMDNDYRRGKLTNHRKVGWETAILAGDALLKAFELISS 128
 126 DDLTDEVIKVLQRLSIAGSHGVGQMLDMQSEGQPIDLETLEMIHKTGTGALLTFA 185
 129 TPNISTEQKALQILQAQGVGQMLDMQSEGQPIDLETLEMIHKTGTGALLTFA 187
 186 VMSAADIAN---VDDTTKEHLESYSYHLGMFQIKDDLLDCYGDGDEAKLGGKVGSLDENNK 242
 188 ALKLGFCPHFTDKRLEQSLQYAEAGLAQVQDDIILDGDSAEIGKQVQADLDLQK 247
 243 STYVSLGKDGABDKLTYHRDAVDELTOIDQSFNTKHLLEIV 285
 248 STYPKLLGSGAKQAQDLQSALESLEKIPEDTTVTRALAEFTI 290
 RESULT 6
 GQPP_CATRO
 ID GQPP_CATRO STANDARD; PRT; 357 AA.
 AC Q42698;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Geranylgeranyl pyrophosphate synthetase, chloroplast precursor (GGPP
 DE synthetase) (GGPS) [includes: Dimethylallyltransferase (EC 2.5.1.1);
 DE Geranyltransferase (EC 2.5.1.10); Farnesyltransferase
 DE (EC 2.5.1.29)].
 GN GGPS1 OR GGCI.
 OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
 OC Vinceae; Catharanthus.
 OX NCBI_TaxID=4058;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. G. Don c20;
 RC Bantignies B., Liboz T., Ambid C.;
 RT "Nucleotide sequence of a Catharanthus roseus geranylgeranyl
 RT pyrophosphate synthase gene";
 RL (in) Plant Gene Register PGR95-119.
 CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
 CC IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.
 CC -1- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
 CC diphosphate = diphosphate + geranyl diphosphate.
 CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
 CC = diphosphate + trans-trans-farnesyl diphosphate.
 CC -1- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
 CC diphosphate = diphosphate + geranylgeranyl diphosphate.
 CC -1- PATHWAY: FIRST COMMITTED STEP IN CAROTENOID BIOSYNTHESIS. KEY
 CC ENZYME IN PLANT TERPENOID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X92893; CAA63486.1;
 DR InterPro; IPR000092; Polyrenyl_synth.
 DR Pfam; PF00348; polyrenyl_synth; 1.
 DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
 DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
 KW transferase; isoprene biosynthesis; Carotenoid biosynthesis;
 KW Chloroplast; Transit peptide.
 FT TRANSIT 1 40
 FT CHAIN 41 357 GERANYLGERANYL PYROPHOSPHATE SYNTHETASE.
 SQ SEQUENCE 357 AA; 38786 MW; 81C52FDEA1E06F8 CRC64;
 Query Match 33.7%; Score 489.5; DB 1; Length 357;
 Best Local Similarity 41.9%; Pred. No. 5.9e-29;
 Matches 111; Conservative 48; Mismatches 97; Indels 9; Gaps 3;
 QY 20 AINKSVMDT-----QLEESMLYSLNAGGKRIRPVLLTLDLSLNTVEYELGMKSAIALE 72
 76 SYNKALEDVAVLREPLKIHESMYSLLAGGKVRPMLCIAACELFGGTSVAMPACAVE 135
 73 MHTVSLIHDDLPAMDNDYRRGKLTNHRKVGWETAILAGDALLKAFELISSDDR-LTD 131
 136 MHTVSLIHDDLPAMDNDYRRGKLTNHRKVGWETAILAGDALLKAFELISSDDR 195
 132 EVKIKVQLRLSIAGSHGVGQMLDMQSEG-QPIDLETLEMIHKTGTGALLTFAVMSAA 190
 196 ERIVRVVGLAKICGSEGLVAGVDVDCSEGIADVGLEHLEFHIHKTAAALGSGSVLGA 255
 191 DIANVDDTTKEHLESYSYHLGMFQIKDDLLDCYGDGDEAKLGGKVGSLDENNKSYVSLG 250
 256 IVGGANDEQISKLRFARCIIGLLFQVVDIILDTVTKSSQELGKTAGDLVADKVTYPKLLG 315
 251 KGAEDKLTYYHRDAVDELTOIDQEQ 275
 316 IDKSREFAEKLNRQAQELAEFDPE 340

```

RESULT 7
ID ISPA_ECOLI STANDARD; PRT: 299 AA.
AC P22939;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Geranyltransferase (EC 2.5.1.10) [Farnesyl-diphosphate synthase]
DE (FPP synthase).
GN ISPA OR B0421.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=91210228; PubMed=2089044;
Fujisaki S., Hara H., Nishimura Y., Horiuchi K., Nishino T.;
"Cloning and nucleotide sequence of the ispA gene responsible for
farnesyl diphosphate synthase activity in Escherichia coli.";
J. Biochem. 108:995-1000(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97426617; PubMed=9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
"The complete genome sequence of Escherichia coli K-12.";
Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
Duncan M., Federspiel N., Hyman R., Kallan S., Komp C., Kurdi O.,
Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
- diphosphate + trans,trans-farnesyl diphosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D00694; BAA00599.1; -
DR EMBL; AF000148; AAC73524.1; -
DR EMBL; U82664; ABA40177.1; -
DR PIR; JQ0665; JQ0665.
DR SWISS-2DPAGE; P22939; COLI.
DR EcoGene; EG10508; ispa.
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; polyprenyl_synth; 1
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR TRANSFERASE; Isoprene biosynthesis; Complete proteome.
SQ SEQUENCE 299 AA; 32159 MW; 15BADD5E135060CA CRC64;
-----
Query Match 33.4%; Score 485.5; DB 1; Length 299;
Best local Similarity 41.8%; Pred. NO. 9.3e-29;
Matches 118; Conservative 42; Mismatches 117; Indels 5; Gaps 4;
QY 10 IDEVNNELSVAINK-SVMDTOLESMYSLNAGGKRIRPVLILLTLDLSINTEYELGMSKA 68
DB 11 VQANQALSRTAPLPFQNTPTVETMQYCALLGGRRLRPFLVYATGHMFGVSTNTLDAPA 70
-----
QY 69 TALEMIHTYSLIHDDLPAWDDYRRGKLTNHNKVTAILAGDALLTKAFELISSDD- 127
|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 71 AAVECIHAYSLTHDDLPAWDDYRRGKLTNHNKVTAILAGDALLTKAFELISSDD- 130
|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 128 -RLTDEVKIKVLRSLIASGHVGMVGMQMDQSGQPIDLETLEMIHKTGKALLTFAV 186
::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 131 PEVSDDRIISMISLASASGIAGCGGQALDLDAGKHPVLPDLALRIHRTGALITRAV 190
|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 187 -MSAADIANVDDTTREHLESYSHLGMFMFQIKDDLLDCYGDRAKLGKVGSDLENNKSY 245
|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 191 RLGALSAGDKGRALPVLDKVAESIGLAFQVQDDILDVVGDTATLGKRGAGDQOLGKSY 250
|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 246 VSLGKGAEDKLTTHRAAANVDDELQIDEQ-FNTHKLLLEIVD 286
|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 251 PALLLEQARKKARDLIDDARQSLKQLAEQSLDTSALEALAD 292
|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
-----
RESULT 8
GGPP_CAPAN STANDARD; PRT: 369 AA.
ID GGPP_CAPAN P80042;
AC P80042;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Geranylgeranyl pyrophosphate synthetase, chloroplast precursor (GGPP
synthetase) (GGPS) [Includes: Dimethylallyltransferase (EC 2.5.1.1);
Geranyltransferase (EC 2.5.1.10); Farnesyltransferase
(EC 2.5.1.29)].
GN GGPS1.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Lamuyo; TISSUE=Fruit;
RX MEDLINE=93272043; PubMed=1303794;
Kuntz M., Roemer S., Sulre C., Huguene P., Weil J.H., Schantz R.,
Camara B.;
"Identification of a cDNA for the plastid-located geranylgeranyl
pyrophosphate synthase from Capsicum annuum: correlative increase in
enzyme activity and transcript level during fruit ripening.";
Plant J. 2:25-34(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Yolo Wonder;
RX MEDLINE=95195169; PubMed=7888631;
Badillo A., Steppuhn J., Deruere J., Camara B., Kuntz M.;
"Structure of a functional geranylgeranyl pyrophosphate synthase gene
from Capsicum annuum.";
Plant Mol. Biol. 27:425-428(1995).
CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.
CC -1- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
diphosphate - diphosphate + geranyl diphosphate.
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
- diphosphate + trans,trans-farnesyl diphosphate.
CC -1- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
diphosphate - diphosphate + geranylgeranyl diphosphate.
CC -1- PATHWAY: FIRST COMMITTED STEP IN CAROTENOID BIOSYNTHESIS. KEY
ENZYME IN PLANT TERPENOID BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----

```


[2]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. Columbia;
RA MEDLINE=20083488; PubMed=106117198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaler B., Mache R., Mueller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Botry M., Bancroft I.,
RA Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dickte W.,
RA Moofjan P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysschaert C., Giesen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
Clark L., Doggett J., Hall S., Kay M., Lennard N., McClay K., Mayes R.,
Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argirion A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefod F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spith J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonolou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana.";
Nature 402:769-777(1999).
-!- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.
-!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
diphosphate = diphosphate + geranyl diphosphate.
-!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
= diphosphate + trans,trans-farnesyl diphosphate.
-!- CATALYTIC ACTIVITY: trans-trans-farnesyl diphosphate + isopentenyl
diphosphate = diphosphate + geranylgeranyl diphosphate.
-!- PATHWAY: FIRST COMMITTED STEP IN CAROTENOID BIOSYNTHESIS. KEY
ENZYME IN PLANT TERPENOID BIOSYNTHESIS.
-!- SUBUNIT: MONOMER.
-!- SUBCELLULAR LOCATION: Chloroplast.
-!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; L25813; AAA32797.1; -

DR EMBL; Z99708; CAB16803.1; -
DR EMBL; AL161590; CAB80347.1; -
DR InterPro; IPR000092; Polyprenyl_synt.
DR Pfam; PF00348; polyprenyl_synt; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW Transferase; Isoprene biosynthesis; Carotenoid biosynthesis;
KW Chloroplast; Transit peptide.
FT TRANSIT 1 ? CHLOROPLAST.
FT CHAIN ? 371 GERANYLGERANYL PYROPHOSPHATE SYNTHETASE.
FT CONFLICT 108 108 R -> S (IN REF. 1).
FT CONFLICT 141 141 A -> R (IN REF. 1).
FT CONFLICT 192 192 A -> S (IN REF. 1).
SQ SEQUENCE 371 AA; 40174 MW; EFA8088A75B6A005 CRC64;

Query Match 32.9%; Score 478; DB 1; Length 371;
Best Local Similarity 40.5%; Pred. No. 4, 4e-28;
Matches 111; Conservative 49; Mismatches 100; Indels 14; Gaps 4;

QY 6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGRIRPVLLLLTLDLSINTYEYELGM 65
Db VNKALDS-----AVPLREPL---KTHEAMRYSLLAGKRVPLVLCIAACELVGGESTAM 138
QY 66 KSATALEMIHYSLIHDDLPAMDNDYRGRKLTNKHVYGEWTAIAGDALLTKAFELI-- 123
Db PAACAVEMIHMTSLIHDDLPAMDNDYRGRKLTNKHVYGEWTAIAGDALLTKAFELI-- 198
QY 124 -SSDRLTDEVKIKVLQRLSIASGHVGMVGGMLDMQSEG---QPIDLETLEMIRKTKTG 179
Db ATSSDVSPVRVRAVAGELAKAIGTEGLVAGQVDISSEGLDNDVGLHEHFIHLHKA 258
QY 180 ALLTFVMSAADIANDVDTTKEHLESYSHYLGMPQIKDDLDYCDGAKLGGKYGSDLE 239
Db ALLEASAVILGAVGGSDDETERLKFKARCIGLLFQVVDLDVTKSKKELGKTAGKDIL 318
QY 240 NKSFTYVSLLLGKDAEDKLTIVRDAVDELQID 273
Db ADKLTYPKIMGLEKRSFAEKLNRREARDOLLGFD 352

RESULT 11
IDSA_METJA STANDARD; PRT; 327 AA.
AC Q58270;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional short chain isoprenyl diphosphate synthase [Includes:
DE Farnesyl pyrophosphate synthetase (EC 2.5.1.1) (FPP synthetase)
DE (Dimethylallyltransferase); Geranyltransferase (EC 2.5.1.10)].
GN IDSA OR M70860.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_Taxid=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=868087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
diphosphate = diphosphate + geranyl diphosphate.
CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate

- diphosphate + trans-farnesyl diphosphate.
- COFACTOR: REQUIRES MAGNESIUM (BY SIMILARITY).
- SUBUNIT: HOMODIMER (BY SIMILARITY).
- SUBCELLULAR LOCATION: CYTOSOL (BY SIMILARITY).
- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U67530; AAB98865.1; --
TIGR; M0860; --
InterPro; IPR000092; Polyprenyl_synt.
Pfam; PF00348; polyprenyl_synt; 1.
PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
Lipid synthesis; Isoprene biosynthesis; Transferrase; Magnesium;
Multifunctional enzyme; Complete proteome.
SEQUENCE 327 AA; 37568 MW; 07D68AC9BD657DAC CRC64;

Query Match 30.1%; Score 438; DB 1; Length 327;
Best Local Similarity 41.2%; Pred. No. 3.2e-25;
Matches 115; Conservative 45; Mismatches 87; Indels 32; Gaps 9;

9 LIDEVNEVNSVAIKSVMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSLNTYEGLMKA 68
17 ILQIDIEELTYVDK---DDKLYNASKHLLFAGGKRIRPVLTVTYMLKKDDIEVLPA 73
69 IALEMHTYSLIHDDLPAMDNDYRGKLTNKHVYGEWTAIAGDALLTKAFELISSDDR 128
74 AAYELHNTYLIHDDI--MDNDERRGKPTVHYVGEPMALAGDILLYAKAFEAVS---R 128
129 LTDEVKI-KVLRSLSTAGSHVGVGMQMLDMQSEG-QPIDLEMLHKTGTGALLTF 186
129 IKDKKAHEVVKILSKAC--VEVCEQAMDMEFYPTWEEYLDIRK-KTGALLLEASV 185
187 MSAADIANVDVDTTKEHLESYHGLMMFQIKDLDLCYGDGDEAKLKGKVGSDLENKSTYV 246
186 GIGAVMADCNDEERALKYKAKRIGLTFQIDVDLIGDQKLGKPGVSDIREGKKTII 245
247 SLGKGDGAEKLYHRDAVDELTDQEQNTKHLEIV 285
246 -----VIHALKTLEED-KKKRLLEIL 265

RESULT 12

ID CRTE_PANAN STANDARD; PRT; 302 AA.
AC P21684;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Geranylgeranyl pyrophosphate synthetase (EC 2.5.1.29) (GGPP synthetase) (Farnesyltransferase).
GN CRTE.
OS Pantoea ananas (Erwinia uredovora).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=553;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20D3;
RX MEDLINE=91072214; PubMed=2254247;
RA Misawa N., Nakagawa M., Kobayashi K., Yamano S., Izawa Y., Nakanura K., Harashina K.;
RT "Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli.";
RL J. Bacteriol. 172:6704-6712(1990).

- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl diphosphate -> diphosphate + geranylgeranyl diphosphate.
- PATHWAY: Carotenoid biosynthesis.
- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D90087; BAA14124.1; --
PIR; A37802; A37802.
InterPro; IPR000092; Polyprenyl_synt.
Pfam; PF00348; polyprenyl_synt; 1.
PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW Carotenoid biosynthesis; Isoprene biosynthesis; Transferrase.
SQ SEQUENCE 302 AA; 32583 MW; CAD804699D2EBA4A CRC64;

Query Match 25.9%; Score 376.5; DB 1; Length 302;
Best Local Similarity 35.6%; Pred. No. 9.5e-21;
Matches 100; Conservative 48; Mismatches 118; Indels 15; Gaps 5;

6 MNKIDDEVNNEVNSVAIKSVMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSL--TEYEL 63
23 IDRLDQL---LPVEGERDVGAAAMEGAL---APGKRIRPMLLLLTARDLGCAVSHDG 75
64 GKMSATALEMIHTYSLIHDDLPAMDNDYRGKLTNKHVYGEWTAIAGDALLTKAFELI 123
76 LLDLACAVEVHVAASLIIDDDPCMDDAKLRRGPTTHSHYGEHVATLAVALLSKAFGVI 135
124 SSDRLTDEVKIKVLRSLSTAGSHVGVGMQMLDMQSEGQPIDLEMLHKTGTGALLT 183
136 ADADGLTPLAKNAVSELSNAIGMQLVQGFQKDLSEGDKPSRAEAILMTNHEKTTLPFC 195
184 FAVMSAADIANVDVDTTKEHLESYHGLMMFQIKDLDLCYGDGDEAKLKGKVGSDLENKNS 243
196 ASMQMASIVANASSEASDCLHFRSLDGOAFQLDLDLTDGMDTGTG-----DSNQDAGKS 250
244 TVVSLGKGDGAEKLYHRDAVDELTDQEQNF-TKHLE 283
251 TLVNLGLPRAVEERLQHLQLASEHLSAACQGHGATQHFQI 291

RESULT 13

ID IDSA_METTM STANDARD; PRT; 324 AA.
AC Q53479;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional short chain isoprenyl diphosphate synthase [includes: Farnesyl pyrophosphate synthetase (EC 2.5.1.1) (FPP synthetase) (dimethylallyltransferase); Geranyltransferase (EC 2.5.1.10)].
GN IDSA.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=79929;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=95070145; PubMed=7979381;
RA Chen A., Poulter C.D.;
RT "Isolation and characterization of idsa: the gene for the short chain isoprenyl diphosphate synthase from Methanobacterium thermoautotrophicum.";
RL Arch. Biochem. Biophys. 314:399-404(1994).
CC - CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl diphosphate -> diphosphate + geranyl diphosphate.
CC - CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: AE000082; AAB91752.1; -
DR InterPro: IPR000092; Polyrenyl_synt.
DR Pfam: PF00348; polyrenyl_synt; 1.
DR PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
DR PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
KW Transferase: Isoprene biosynthesis; Plasmid.
SQ SEQUENCE 332 AA; 34688 MW; EE68C154749AA87 CRC64;

Query Match 24.2%; Score 352; DB 1; Length 332;
Best Local Similarity 36.5%; Pred. No. 6.7e-19;
Matches 92; Conservative 40; Mismatches 108; Indels 12; Gaps 3;
QY 27 DTQLESMYSLNAGGRIRPVLLLTLDLSLNTYEYELGKMSATALEMIHTYSLIHDDLPA 86
59 ETELMAAMRYATLHGGRTRALLCLAAAGALADTPAHILDDVGAAIEMMHACTLVHDDLPA 118
QY 87 MDNDVYRRGKLTNKHVYGEWTAILAGDALTKAFELISSDDRLTDEVKIKVLQRLSTASG 146
Db 119 MDDVLRGLPTVHVYFEGENTALLVGDALQAHAFLLASLDAPGDN-RIALVRELQAQVS 177
QY 147 HVGWVGQMLDMQSEGPIDLETLEMIHKTGKTGALLTFAYMSAADIANVDD-----TTKEH 202
Db 178 AEGAAGQAMDLSLVGKHVELDRIVAMHRMKCGALVRASVRMGALCAIAEDAADATLYCA 237
QY 203 LESYSHLGMFMQIKDDLLDCYGEAKLGGKVGSDLENNKSTVYSLGKDGAEKLTYYHR 262
Db 238 LDHYSACFGLAQVDDILDATATLGTGKTPGKDAAAOKPTCASIMGLQAA-----R 290
QY 263 DAAVDELTOIDE 274
Db 291 QFALDLLCEAGE 302

Search completed: May 29, 2003, 11:59:11
Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:54:20 ; Search time 33 Seconds
(without alignments)
1791.986 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPNKLIDEVNNELSA.....ELTQIDQFNKHLLEIVDL 287

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phage:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_virus:**
- 16: sp_bacteriaph:**
- 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1444	99.4	293	16 Q99TX2	Q99TX2 staphylococ
2	702.5	48.3	293	16 Q8Y7C2	Q8Y7C2 listeria mo
3	686.5	47.2	293	16 Q92BZ1	Q92BZ1 listeria in
4	638	43.9	294	16 Q9K969	Q9K969 bacillus ha
5	576	39.6	297	16 Q8RE07	Q8RE07 fusobacteri
6	551.5	38.0	290	16 Q99YX5	Q99YX5 streptococ
7	551	37.9	301	16 Q8XJE0	Q8XJE0 clostridium
8	549	37.8	291	16 Q97QK0	Q97QK0 streptococ
9	540.5	37.2	285	16 Q9CH81	Q9CH81 lactococcus
10	535.5	36.9	383	10 Q8WLR9	Q8WLR9 abies grand
11	534.5	36.8	309	16 Q8Z085	Q8Z085 anaerobac
12	534	36.8	307	2 Q9S5F1	Q9S5F1 synechococ
13	532	36.6	289	16 Q97HD2	Q97HD2 clostridium
14	526.5	36.2	302	16 P72683	P72683 synechocyst
15	520.5	35.8	393	10 Q9ZPM3	Q9ZPM3 taxus canad
16	519	35.7	367	10 Q9SXZ6	Q9SXZ6 daucus caro

17	516.5	35.5	294	16 Q8RAC7	Q8RAC7 thermoanaer
18	516.5	35.5	366	10 Q9SXZ5	Q9SXZ5 daucus caro
19	499.5	34.4	356	10 Q81099	Q81099 helianthus
20	489	33.7	361	10 Q9FV47	Q9FV47 tagetes ere
21	485.5	33.4	298	16 Q9JSM0	Q9JSM0 neisseria m
22	485.5	33.4	299	16 Q8XE75	Q8XE75 escherichia
23	484	33.3	294	16 Q9KTL2	Q9KTL2 vibrio chol
24	483.5	33.3	288	10 Q94IF0	Q94IF0 eucommia ul
25	482.5	33.2	370	10 Q94ID7	Q94ID7 hevea brasili
26	481	33.1	295	16 Q9CNA1	Q9CNA1 pasteurella
27	480.5	33.1	298	2 Q9RME5	Q9RME5 zymomonas m
28	478.5	32.9	316	10 Q42866	Q42866 lupinus alb
29	478.5	32.9	368	10 Q9SSU0	Q9SSU0 croton subl
30	475.5	32.7	306	16 Q8ZC46	Q8ZC46 pseudomonas
31	474.5	32.7	295	16 Q9HWY4	Q9HWY4 pseudomonas
32	473	32.6	299	16 Q8ZRD0	Q8ZRD0 salmonella
33	473	32.6	299	16 Q8Z8X2	Q8Z8X2 salmonella
34	473	32.6	360	10 Q9LUE1	Q9LUE1 arabidopsis
35	470.5	32.4	294	16 Q9A6M4	Q9A6M4 caulobacter
36	469.5	32.3	259	16 Q9K1A6	Q9K1A6 neisseria m
37	469	32.3	376	10 Q9XHX1	Q9XHX1 oryza sativ
38	467	32.1	347	10 Q9ZU77	Q9ZU77 arabidopsis
39	465.5	32.0	377	10 Q9SBR3	Q9SBR3 mentha pipe
40	462	31.8	299	16 Q8XX94	Q8XX94 ralstonia s
41	462	31.8	360	10 Q9LUD9	Q9LUD9 arabidopsis
42	457.5	31.5	335	16 Q8UBX7	Q8UBX7 agrobacteri
43	456	31.4	360	10 Q9LHRA	Q9LHRA arabidopsis
44	452	31.1	357	10 Q9LIA0	Q9LIA0 arabidopsis
45	450.5	31.0	306	10 Q9LRR0	Q9LRR0 arabidopsis

ALIGNMENTS

RESULT 1

Q99TX2					
ID	Q99TX2	PRELIMINARY;	PRT;	293 AA.	
AC	Q99TX2				
DT	01-JUN-2001 (TREMREL. 17, Created)				
DT	01-JUN-2001 (TREMREL. 17, Last sequence update)				
DT	01-MAR-2002 (TREMREL. 20, Last annotation update)				
DE	IsPa protein (Geranyltransferase homolog).				
GN	ISPA OR SAV1521 OR SA1352.				
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699), and				
OS	Staphylococcus aureus (strain N315).				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;				
OC	Staphylococcus.				
OX	NCBI_TaxID=158878, 158879;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);				
RX	MEDLINE=21311952; PubMed=11418146;				
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,				
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,				
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,				
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,				
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,				
RA	Kanehisa M., Masahita A., Oshima K., Furuya K., Yoshino C., Shiba T.,				
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,				
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus				
RL	aureus";				
RL	Lancet 357:1225-1240(2001).				
DR	EMBL; AP003362; BAB57683.1; -				
DR	EMBL; AP003134; BAB42614.1; -				
DR	InterPro; IPR000092; Polyprenyl_synth.				
DR	Pfam; PF00348; polyprenyl_synth. 1				
DR	PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.				
DR	PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.				
KW	Transferrase; Complete proteome.				
SQ	SEQUENCE 293 AA; 32746 MW; E7AFFBBA868F7D8 CRC64;				
Query Match		99.4%;	Score 1444;	DB 16;	Length 293;
Best Local Similarity		99.3%;	Pred. No. 8.6e-103;		

Matches 285; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTNLPNMKLLDEVNLSVAINKSVMDTQLEESMLYSINAGGKRIRPVLLLLTDSLNTE 60
 DB 1 MTNLPNMKLLDEVNLSVAINKSVMDTQLEESMLYSINAGGKRIRPVLLLLTDSLNTE 60
 QY 61 YELGMSKSAIALEMHTYSLIHDDLPAMDNDYRGLTNHVKYGEWTAIAGDALLTKAF 120
 DB 61 YELGMSKSAIALEMHTYSLIHDDLPAMDNDYRGLTNHVKYGEWTAIAGDALLTKAF 120
 QY 121 ELISSDRLTDEVKIKVLRSLASGHVGMVGMQMDQSEGQIDLETLEMHKTKTGA 180
 DB 121 ELISSDRLTDEVKIKVLRSLASGHVGMVGMQMDQSEGQIDLETLEMHKTKTGA 180
 QY 181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMFMFQIKDDLLDCYGDGAKLGKVGSDLEN 240
 DB 181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMFMFQIKDDLLDCYGDGAKLGKVGSDLEN 240
 QY 241 NKSTYVSLGKGAEDKLTHRDAADVDELQIDQEFNTHLLEIVDL 287
 DB 241 NKSTYVSLGKGAEDKLTHRDAADVDELQIDQEFNTHLLEIVDL 287

RESULT 2

Q8Y7C2 PRELIMINARY; PRT; 293 AA.

AC Q8Y7C2
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical protein lmo1363.
 GN LMO1363.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RC STRAIN=EGD-E / SEROVAR 1/2A;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Rammel B., Rose M., Schluter T., Simoes N., Trierz A.,
 Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of *Listeria species*."
 RL Science 294:849-852(2001).
 DR EMBL: AL591978; CAC9441.1; -
 DR Listlist: LMO1363;
 DR InterPro: IPR000092; Polyprenyl_synth.
 DR InterPro: IPR000235; Ribosomal_S7.
 DR Pfam: PF00348; polyprenyl_synth; 1.
 DR PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
 DR PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
 DR PROSITE: PS00052; RIBOSOMAL_S7; UNKNOWN_1.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 293 AA; 32308 MW; 6A7F0C75B1824CC9 CRC64;

Query Match 48.3%; Score 702.5; DB 16; Length 293;
 Best Local Similarity 52.9%; Pred. No. 5.6e-46;
 Matches 148; Conservative 46; Mismatches 81; Indels 5; Gaps 2;

QY 8 KLIDEVNNLSVAINKSVMDTQLEESMLYSINAGGKRIRPVLLLLTDSLNTEYELGMS 67
 DB 13 KVIDE---SLFKEINERNIEPRKESMLYSIQAGGKRIRPMLVFATLQALKVNLGVKT 69
 QY 68 AIALEMIHTYSLIHDDLPAMDNDYRGLTNHVKYGEWTAIAGDALLTKAFELISSDD 127

DB 70 ATALEMHTYSLIHDDLPAMDNDYRGLTNHVKYGEWTAIAGDALLTKAFELISSDD 129
 QY 128 RLTDVKKIKVLRSLASGHVGMVGMQMDQSEGQIDLETLEMHKTKTGAITFAVM 187
 DB 130 NLSFETRIALINQISVSSGAEVGMVGMQMDQSEGQIDLETLEMHKTKTGAITFAVM 189
 QY 188 SAADIANVDDTTKEHLESYSYHLGMFMFQIKDDLLDCYGDGAKLGKVGSDLENKSTYVS 247
 DB 190 SAAKIAEADPEQTKRLRIFAENIGIGFQISDDILDVIGDETKMGKTGADAFLNKSTYPS 249
 QY 248 LLGKGAEDKLTHRDAADVDELQIDQEFNTHLLEIVDL 287
 DB 250 LLTLDGAKRALNEHVITAKSALSGHD--FDDEILLKLADL 287

RESULT 3

Q92BZ1 PRELIMINARY; PRT; 293 AA.

AC Q92BZ1
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Hypothetical protein lni1400.
 GN Lni1400.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Rammel B., Rose M., Schluter T., Simoes N., Trierz A.,
 Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of *Listeria species*."
 RL Science 294:849-852(2001).
 DR EMBL: AL596168; CAC96631.1; -
 DR Listlist: LNI01400;
 DR InterPro: IPR000092; Polyprenyl_synth.
 DR InterPro: IPR000235; Ribosomal_S7.
 DR Pfam: PF00348; polyprenyl_synth; 1.
 DR PROSITE: PS00723; POLYPRENYL_SYNTHET_1; UNKNOWN_1.
 DR PROSITE: PS00444; POLYPRENYL_SYNTHET_2; UNKNOWN_1.
 DR PROSITE: PS00052; RIBOSOMAL_S7; UNKNOWN_1.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 293 AA; 32459 MW; 36649E7E47D8988D CRC64;

Query Match 47.2%; Score 686.5; DB 16; Length 293;
 Best Local Similarity 52.1%; Pred. No. 9.5e-45;
 Matches 146; Conservative 45; Mismatches 84; Indels 5; Gaps 2;

QY 8 KLIDEVNNLSVAINKSVMDTQLEESMLYSINAGGKRIRPVLLLLTDSLNTEYELGMS 67
 DB 13 KVIDE---SLFKEINERNIEPRKESMLYSIQAGGKRIRPMLVFATLQALNIEMRGLKT 69
 QY 68 AIALEMIHTYSLIHDDLPAMDNDYRGLTNHVKYGEWTAIAGDALLTKAFELISSDD 127
 DB 70 ATALEMHTYSLIHDDLPAMDNDYRGLTNHVKYGEWTAIAGDALLTKAFELISSDD 129
 QY 128 RLTDVKKIKVLRSLASGHVGMVGMQMDQSEGQIDLETLEMHKTKTGAITFAVM 187
 DB 130 NLSFETRIALINQISVSSGAEVGMVGMQMDQSEGQIDLETLEMHKTKTGAITFAVM 189

```
QY 188 SAADIANVDDTTKHELESYSYHLGMMFOIKDLDLCYGDCAKKGKVGSDLENNKSTYVS 247
DB 190 SAAKTAETPTQKRLRIFAENIGIGFQISDILIDVIGDETGMGKRGKGVDAFLNKRSTYPG 249
QY 248 LIGKDGAEKLTYYHDAVDELQIDQFNTHKLEIVDL 287
DB 250 LITLEGAKRALNEHVSIAKSALSGHD--FDDEILLKLADL 287

RESULT 4
ID Q9K969 PRELIMINARY; PRT; 294 AA.
AC Q9K969;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Geranyltransferase.
GN BH2781.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
DR EMBL; AP001516; BAB06500.1; -.
DR InterPro; IPR000092; PolyPrenyl_synth.
DR Pfam; PF00348; polyPrenyl_synth.1;
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 294 AA; 32633 MW; 3B2942868A62AD9E CRC64;

Query Match 43.9%; Score 638; DB 16; Length 294;
Best Local Similarity 47.7%; Pred. No. 5e-41;
Matches 134; Conservative 53; Mismatches 92; Indels 2; Gaps 1;

QY 6 MNKLDVNNELSVAINKSVMDTOLESMLYSLNAGGRIRPVLLLLTDSINTYEYELGM 65
DB 9 LDEIKDITEERPAHIERLNSPDMKNSMLYSLKAGGRIRPALLATMKSFQKDISQGI 68
QY 66 KSAIALEMTHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAILAGDALLTKAFELISS 125
DB 69 DLACAIEMLHTYSLIHDDLPMSDDDDIRGKPTNKHVGEAAILAGDALLTYSFEIVAK 128
QY 126 DDLRLTEVKIKVLRSLASGHVGMVGMQMDQSEGQIDLETLEMIHKTGTGALLTFA 185
DB 129 MKGVDPARTLCIEELARAAGPEGVMGVQVADIEGKNQKLTVEGLEIYHHHTGALLSPA 188
QY 186 VMSADIANVDDTTKHELESYSYHLGMMFOIKDLDLCYGDCAKKGKVGSDLENNKSTY 245
DB 189 IVAGARLADASEQDIENIRFRSRELGLLFQIKDDILDVEGQQAIGKPVGSDGNGKSTY 248
QY 246 VSLCKDGAEKLTYYHDAVDELQIDQFNTHKLEIVD 286
DB 249 FSLTLEGAKELHLHTLLAKELYLSV--QMNRLLLELTD 287

RESULT 5
ID Q8RE07 PRELIMINARY; PRT; 297 AA.
AC Q8RE07;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Dimethylallyltransferase (EC 2.5.1.1).
```

```
GN FN1327.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteriia; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasleva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Ponstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018 (2002).
DR EMBL; AE010637; AAL95523.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 297 AA; 33577 MW; 89DC35727945FC53 CRC64;

Query Match 39.6%; Score 576; DB 16; Length 297;
Best Local Similarity 44.2%; Pred. No. 2.8e-36;
Matches 123; Conservative 54; Mismatches 99; Indels 2; Gaps 1;

QY 11 DEVNNELSVAINKSVMDTOLESMLYSLNAGGRIRPVLLLLTDSINTYEYELGMKSAIA 70
DB 13 DFEFTELKKEKELSYETIAKMEYAILNGKRLRPFLFVTLLENININKGVKSAIA 72
QY 71 LEMITHTYSLIHDDLPAMDNDYRRGKLTNKHVYGEWTAILAGDALLTKAFELISSD--DR 128
DB 73 LEMIHYSYLVHDDLPALDNDYRRGKLTNKHVYGEAAILAGDALLTAFYVLSQNKLEL 132
QY 129 LTDEVKIKVLRSLASGHVGMVGMQMDQSEGQIDLETLEMIHKTGTGALLTFAVMS 188
DB 133 LSSKQIVNIISKTSEYAGIDGMIGGQIMIDQSENKKIDLETKYIHSHTGKLIKLP 192
QY 189 AADIANVDDTTKHELESYSYHLGMMFOIKDLDLCYGDCAKKGKVGSDLENNKSTYVSL 248
DB 193 ACIIANLEKDKREVLVEYADLIGAFQVKKDDILDVEGTFEDLGKPVGSDVHLKATPSI 252
QY 249 LGKDGAEKLTYYHDAVDELQIDQFNTHKLEIVD 286
DB 253 LGMESEKKTLLNTVEKAKELIKNKFGEEKGVLLISLAD 290

RESULT 6
ID Q99YX5 PRELIMINARY; PRT; 290 AA.
AC Q99YX5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative geranyltransferase (farnesyl diphosphate synthase)
DE (EC 2.5.1.10).
GN FPS OR SPY1498.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvarov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
DR EMBL; AE006583; AAK34297.1; -.
DR InterPro; IPR000092; PolyPrenyl_synth.
DR Pfam; PF00348; polyPrenyl_synth.1;
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
```

```
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 290 AA; 31824 MW; 875FA2DAE53A0331 CRC64;

Query Match 38.08; Score 551.5; DB 16; Length 290;
Best Local Similarity 44.8%; Pred. No. 2.1e-34;
Matches 130; Conservative 46; Mismatches 99; Indels 15; Gaps 6;

QY 6 MNKL--IDEVNNELSVAINSVMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSL-----N 58
DB 1 MDKLARIDEAIRRYKTTNGVSE-ELIDAILYSVDSGGKRIRPILILEMIEFGVSLQN 59
QY 59 TEYELGKMSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHNKHYGEWTAILAGDALLTK 118
DB 60 AHFDL----AAALEMIHTGSLIHDDLPAMDNDYRRGKLTNHNKHYGEWTAILAGDLSFLD 115
QY 119 AFEISSDDRLTDEVKIKVLQRLSIASGHVGMVGOMLQMSGQPIDLETLEMIHKT 178
DB 116 PFLGIAQAE--LNSEVKVALLQELSLASGTFTGMVGOMLQMSGQPIDLETLEMIHKT 174
QY 179 GALLTFVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLDLCYGDCAKLGKVGSD 238
DB 175 GKLLAPFKAALITQEAQMTVROQLAQAGMLIGHAFQIRDDILDVTASFEDLGKTPKDL 234
QY 239 ENNKSYVSLGKGDGAEDKLTYYHRDA--VDELTQIDEQNTKHLLEIVD 286
DB 235 FAEKATYPSLLGLEASYQLLTESLIDQALTIFQLESVDGFKPQIITKLE 284

RESULT 7
Q8XJEO PRELIMINARY; PRT; 301 AA.
ID Q8XJEO
AC Q8XJEO;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DE Geranyltransferase.
GN CPE1820.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RA "Complete genome sequence of Clostridium perfringens, an anaerobic
RA flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003191; BAB81526.1; -
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; polyprenyl_synth; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 301 AA; 33760 MW; 3BF349AEF836461 CRC64;

Query Match 37.9%; Score 551; DB 16; Length 301;
Best Local Similarity 41.8%; Pred. No. 2.4e-34;
Matches 118; Conservative 60; Mismatches 100; Indels 4; Gaps 3;

QY 2 TNLPMNKLIDEVNNELSVAINKS-VMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSLNT 60
DB 10 TANNINSLKEEDVQSLKAFYKNDREYNKVLDSMAYSVNGGKRIRPILMLLSYYIKSD 69
QY 61 YELGKMSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHNKHYGEWTAILAGDALLTKAF 120
DB 70 YKLLTPMAAIEMIHTYSLIHDDLPAMDNDYRRGKLTNHNKHYGEWTAILAGDALLNEAM 129
QY 121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGOMLQMSGQPIDLETLEMIHKT 179
DB 121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGOMLQMSGQPIDLETLEMIHKT 179

DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 290 AA; 31824 MW; 875FA2DAE53A0331 CRC64;

Query Match 38.08; Score 551.5; DB 16; Length 290;
Best Local Similarity 44.8%; Pred. No. 2.1e-34;
Matches 130; Conservative 46; Mismatches 99; Indels 15; Gaps 6;

QY 6 MNKL--IDEVNNELSVAINSVMDTQLEESMLYSLNAGGKRIRPVLLLTLDLSL-----N 58
DB 1 MDKLARIDEAIRRYKTTNGVSE-ELIDAILYSVDSGGKRIRPILILEMIEFGVSLQN 59
QY 59 TEYELGKMSAIALEMIHTYSLIHDDLPAMDNDYRRGKLTNHNKHYGEWTAILAGDALLTK 118
DB 60 AHFDL----AAALEMIHTGSLIHDDLPAMDNDYRRGKLTNHNKHYGEWTAILAGDLSFLD 115
QY 119 AFEISSDDRLTDEVKIKVLQRLSIASGHVGMVGOMLQMSGQPIDLETLEMIHKT 178
DB 116 PFLGIAQAE--LNSEVKVALLQELSLASGTFTGMVGOMLQMSGQPIDLETLEMIHKT 174
QY 179 GALLTFVMSAADIANVDDTTKEHLESYSYHLGMFQIKDDLDLCYGDCAKLGKVGSD 238
DB 175 GKLLAPFKAALITQEAQMTVROQLAQAGMLIGHAFQIRDDILDVTASFEDLGKTPKDL 234
QY 239 ENNKSYVSLGKGDGAEDKLTYYHRDA--VDELTQIDEQNTKHLLEIVD 286
DB 235 FAEKATYPSLLGLEASYQLLTESLIDQALTIFQLESVDGFKPQIITKLE 284

RESULT 8
Q97QKO PRELIMINARY; PRT; 291 AA.
ID Q97QKO
AC Q97QKO;
DT 01-OCT-2001 (T-EMBLrel. 18, Created)
DT 01-OCT-2001 (T-EMBLrel. 18, Last sequence update)
DE Geranyltransferase.
GN SP1205.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapfel E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angluoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RA "Complete genome sequence of a virulent isolate of Streptococcus
RA pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL; AE007421; AAK75312.1; -
DR TIGR; SP1205; -
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; Polyprenyl_synth; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 291 AA; 31868 MW; CF372A7642934062 CRC64;

Query Match 37.8%; Score 549; DB 16; Length 291;
Best Local Similarity 46.2%; Pred. No. 3.2e-34;
Matches 121; Conservative 50; Mismatches 87; Indels 4; Gaps 3;

QY 28 TOLEESMLYSLNAGGKRIRPVLLLTLDLSL-TEYELGKMSAIALEMIHTYSLIHDDLP 86
DB 25 SSRESVLYSIHAGGKRIRPFLLEVLQVTKPAHAQVATALEMIHTYSLIHDDLP 84
QY 87 MNDYRRGKLTNHNKHYGEWTAILAGDALLTKAFELTSSDDRLTDEVKIKVLQRLSIASG 146
DB 85 MDDDDYRRGRLTNHKKFGEAMAILAGDALFLDSYALLAQAD-LPSQIKVDLIANLSL 143
QY 147 HVGWVGOMLQMSGQPIDLETLEMIHKTGALLTFVMSAADIANVDDTTKEHLESY 206
DB 144 SLGMAQVLDMEGEHQHLSLEELQTHANKTKLLAYPFAAAITAEISPENQVCLKTV 203
QY 207 SYHLGMFQIKDDLDLCYGDCAKLGKVGSDLENKSYVSLGKGDGAEDKLTYYHRDAV 266
DB 204 GELIGAFQVRDDVLDVDTASFEEIGTKPQDKQAERSTYPALLGLEESIAFCNOTLDQ 263
QY 267 DELTQIDEQ--FNTHKHLLEIVD 286
DB 264 EKLEIAQQLPFETESIVSVE 285
```


RESULT 9

Q9CH81 PRELIMINARY; PRT: 285 AA.
AC Q9CH81;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Farnesyl diphosphate synthase (EC 2.5.1.10).
GN ISPA OR LL0857.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Rolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
lactis ssp. lactis IL1403".
DR EMBL: AE006320; AAK04955.1; -
DR InterPro: IPR000092; Polyprenyl_synth.
DR Pfam: PF00348; polyprenyl_synth.1.
DR PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Complete proteome.
SQ SEQUENCE 285 AA; 31354 MW; 8678F7A51B850C1F CRC64;

Query Match 37.2%; Score 540.5; DB 16; Length 285;
Best Local Similarity 44.8%; Pred. No. 1.4e-33;
Matches 130; Conservative 42; Mismatches 97; Indels 21; Gaps 5;
QY 6 MNKLENNELSVAINKSVMDTOLEESMLYNAGGKRIRPVLLLLTLDLSLNT- 60
DB 1 MDTKILKLEDFTEYESAEPTGLAESAKSYLLAGGKRIRPLFLNLEAFDELKSAH 60
QY 61 YELGMSAIALEMHTYSLIHDDLPAMDNDYRRGKLNHKKVYGEWTAILAGDALLTAF 120
DB 61 YHV-----AAALEMHTGSLIHDDLPAMDNDYRRGKLNHKKFDEATAILAGDTLFFDPF 116
QY 121 ELISSDDRRLTDEVKIKVLRSLASGSHVGMVGOMLDMQSEGOPTDLETMHKTGGA 180
DB 117 FILSTAD-LSAEIIVALTRELAFASGSGWAGQILDWAGEKELTAEIQTQIKRKTGR 175
QY 181 LITFAVMSAADIANVDDTTKEHLESYSYHLGMMFOIKDDLLDCYGDGAKLGGKVGSDLEN 240
DB 176 LITPEVAAGIVAOKSTDEVEKLRQVGQILGLAFQIRDDILDVATFAELGKTPGKDILE 235
241 NKSTYVSLLDGAGDEKLTTH-----RDAAV---DELQIDPQFWTK 279
DB 236 EKSTYVAHLGLEGAKSLTGNLSEVKLLTLDLSVTDSEIFKRIIQLQEVK 285

RESULT 10

Q8W1R9 PRELIMINARY; PRT: 383 AA.
AC Q8W1R9;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Geranylgeranyl diphosphate synthase.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=46611;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21671398; PubMed=11733504;
RA Burke C., Croteau R.;
RT "Interaction with the Small Subunit of Geranyl Diphosphate Synthase
Modifies the Chain Length Specificity of Geranylgeranyl Diphosphate

RT Synthase to Produce Geranyl Diphosphate.";
RL J. Biol. Chem. 277:3141-3149(2002).
DR EMBL: AF425235; AAL17614.2; -
DR InterPro: IPR000092; Polyprenyl_synth.
DR Pfam: PF00348; polyprenyl_synth.1.
DR PROSITE: PS00723; POLYPRENYL_SYNTHET_1; UNKNOWN_1.
DR PROSITE: PS00444; POLYPRENYL_SYNTHET_2; UNKNOWN_1.
SQ SEQUENCE 383 AA; 42069 MW; EA8D3A015B51CD31 CRC64;
Query Match 36.9%; Score 535.5; DB 10; Length 383;
Best Local Similarity 44.9%; Pred. No. 5.1e-33;
Matches 122; Conservative 48; Mismatches 83; Indels 19; Gaps 6;
QY 2 TNLP---MNKLD---EVNNELSVAINKSV-----MDTOLEESMLYNAGGKRIRP 47
DB 76 TNLPEKVKKEVFDFKVEYLRSKAMAVNEALDRAVPLRYPERIHEAMRYSLLAGGKRVRP 135
QY 48 VLLLTLDLSLNTYEELGKMSAIALEMHTYSLIHDDLPAMDNDYRRGKLNHKKVYGEWT 107
DB 136 VLCISACELVGQTEEVAMPTACAMEHTMSLIHDDLPAMDNDYRRGKLNHKKVYGEWT 195
QY 108 AILAGDALLTAFELI---SSDRLTDEVKIKVLRSLASGSHVGMVGOMLDMQSEGO- 163
DB 196 AILAGDALLTAFELI---SSDRLTDEVKIKVLRSLASGSHVGMVGOMLDMQSEGO- 254
QY 164 PIDLETLEMIHKTGTCALLTFVMSAADIANVDDTTKEHLESYSYHLGMMFOIKDDLLDC 223
DB 255 SVDLDTLEWIHIHKTAVLLECSVMCGAIISSGASDNEIERIQYARSVGLLFQVDDILDV 314
QY 224 YGDEAKLGGKVGSDLENKSTYVSLLGKDGAE 255
DB 315 TKSSKELGKTAGKDLISDKATYPKLMGLEKAK 346

RESULT 11

Q8Z085 PRELIMINARY; PRT: 309 AA.
ID Q8Z085
AC Q8Z085;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Geranylgeranyl diphosphate synthase.
GN ALR0213;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120".
RL DNA Res. 8:203-213(2001).
DR EMBL: AP003581; BAB77737.1; -
DR InterPro: IPR000092; Polyprenyl_synth.
DR Pfam: PF00348; polyprenyl_synth.1.
DR PROSITE: PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE: PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Complete proteome.
SQ SEQUENCE 309 AA; 33706 MW; E5E2DA7F1E71B657 CRC64;
Query Match 36.8%; Score 534.5; DB 16; Length 309;
Best Local Similarity 47.8%; Pred. No. 4.5e-33;
Matches 119; Conservative 35; Mismatches 92; Indels 3; Gaps 2;
QY 29 QLEESMLYNAGGKRIRPVLLLLTLDLSLNTYEELGKMSAIALEMHTYSLIHDDLPAMD 88
DB 44 KIYESMRYSLLAGGKRIRPLILCLATSEMMGGTIEIATPACAVEMHTMSLIHDDLPAMD 103

```
QY 89 NDDYRRGKLTNHHKVGWETAILAGDALLTKAFELIS--SDRLTDEVKIKVQLRLSTASG 146
Db 104 NDDYRRGKLTNHHKVGWETAILAGDGLLAYAFVAITPESVPRDRVLQWVARGALG 163
QY 147 HVGWVGOMLDMQSEGO-PIDLETLEMIHKTGTGALLTFAVMSAADTANVDDTTKEHLES 205
Db 164 AAGLVGGQVVDLQSEKSDTSLEFLNFHNNHTAALLACVCGGIIAGASSENQVRLSR 223
QY 206 YSYHLGMMFQIKDDLLCYGDEAKRKGKVGSDLENNKSTVSLVGLKDGAEKLTTHRDAA 265
Db 224 YSQNIGLAFQIIDDILTITQEQGLKGTACKDALLAKKVTYPSLWGLWIEQSRVKAQOLLEA 283
QY 266 VDELTQIDE 274
Db 284 CTELEPGE 292

RESULT 12
Q9S5F1 PRELIMINARY; PRT; 307 AA.
Q9S5F1;
01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Geranylgeranyl diphosphate synthase (SeigGRS).
GN CRTE.
OS Synecococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9939254; PubMed=10412909;
RA Ohto C., Ishida C., Nakane H., Muramatsu M., Nishino T., Obata S.;
RT "A thermophilic cyanobacterium Synecococcus elongatus has three
RT different Class I prenyltransferase genes.";
RL Plant Mol. Biol. 40:307-321(1999).
DR EMBL: AB016093; BAA82613.1; -.
DR InterPro: IPR000092; Polyprenyl-synt.
DR Pfam: PF00348; polyprenyl-synt; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
SQ SEQUENCE 307 AA; A3F918D38B7FB648 CRC64;

Query Match 36.8%; Score 534; DB 2; Length 307;
Best Local Similarity 50.9%; Pred. No. 4.9e-33;
Matches 114; Conservative 32; Mismatches 76; Indels 2; Gaps 2;

29 QLEESMLYSLNAGGKRIRPVLLLLTLDLSLNTVEYELGKMSAIALEMIHTVSLIHDDLPA 88
43 KIYDARYSLMAGCKRLRPILCATCELMGGTVEMAMPTACALEMIHTVSLIHDDLPA 102

QY 89 NDDYRRGKLTNHHKVGWETAILAGDALLTKAFELIS--SDRLTDEVKIKVQLRLSTASGH 147
Db 103 NDDYRRGKLTNHHKVGWETAILAGDGLLAYAFVAITPESVPRDRVLQWVARGALG 162
QY 148 VGVWVGOMLDMQSEGO-PIDLETLEMIHKTGTGALLTFAVMSAADTANVDDTTKEHLES 206
Db 163 TGLVGGQVVDLQSEKSDTSLEFLNFHNNHTAALLACVCGGIIAGASSENQVRLSR 222
QY 207 YSYHLGMMFQIKDDLLCYGDEAKRKGKVGSDLENNKSTVSLVGLKDGAEKLTTHRDAA 265
Db 223 AANIGLAFQIIDDILTITQEQGLKGTACKDALLAKKVTYPSLWGLWIEQSRVKAQOLLEA 283

RESULT 13
Q97HD2 PRELIMINARY; PRT; 289 AA.
Q97HD2;
01-OCT-2001 (TEMBLrel. 18, Created)
DT 01-OCT-2001 (TEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Predicted geranylgeranyl pyrophosphate synthase.
GN CAC2080.
```

```
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Berton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007710; AAK80039.1; -.
DR InterPro: IPR000092; Polyprenyl-synt.
DR Pfam: PF00348; polyprenyl-synt; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Complete proteome.
SQ SEQUENCE 289 AA; 32381 MW; C17BAF2903F04DC2 CRC64;

Query Match 36.6%; Score 532; DB 16; Length 289;
Best Local Similarity 44.9%; Pred. No. 6.4e-33;
Matches 127; Conservative 44; Mismatches 108; Indels 4; Gaps 3;

QY 7 NKLI-DEVNNELSVAI-NKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDLSLNTVEYELG 64
Db 3 NKVIKKEVEYLSRYEGKDNYNKRVYESMNYSLNAGGKRIRPVLLLLTLDLSLNTVEYELG 62
QY 65 MKSAIALEMIHTVSLIHDDLPA 124
Db 63 IDIAAIAEMHTVSLIHDDLPA 122
QY 125 SDDRLEDEVKIKVQLRLSTASGHVGVGGOMLDMQSEGO-PIDLETLEMIHKTGTGALLT 184
Db 123 KYCIGGEEALKACLMISKAASDSGDMGGVGVVILSEGKKNEDELRYMHKKTGELIKA 182
QY 185 AVMSAADIANVDDTTKEHLESYVHLGMMFQIKDDLLCYGDEAKRKGKVGSDLENNKST 244
Db 183 AVVSGAILGGAPLHVELLSQYDGKGLAFQIEDDILDIGDKIMKTKSLENDKCT 242
QY 245 YVSLGKDGAEKLTTHRDAAVDELTDQIDQFNTHLLEIVDL 287
Db 243 YVTLYGIDCKICKRELTDCLDIIGKI--QGNTLLEKTEL 283

RESULT 14
P72683 PRELIMINARY; PRT; 302 AA.
ID P72683;
AC P72683;
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Geranylgeranyl pyrophosphate synthase.
GN CRTE OR SLR0739.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90899; BAA16690.1; -.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 19:36:06 ; Search time 2818 Seconds
(without alignments)
2963.983 Million cell updates/sec

Title: US-09-925-637-64

Perfect score: 1453

Sequence: 1 MTNLPKMLIDEVNNELSV.....ELTQIDQFNKHLLEIVDL 287

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

arched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/uspto_pool/US09925637/runat_23052003_174817_23423/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637 -CGN_1_2496 -runat_23052003_174817_23423 -NCPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.pat:*
23: em.ph:*
24: em.pl:*
25: em.ro:*
26: em.sts:*
27: em.un:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	1448	99.7	33750	1	AP004827	AP004827 Staphyloc
C 2	1444	99.4	30350	1	AP003134	AP003134 Staphyloc
C 3	1444	99.4	346900	1	AP003362	AP003362 Staphyloc
C 4	1436	98.8	882	6	AR106466	AR106466 Sequence
5	1015	69.9	3444	1	AF270104	AF270104 Staphyloc
6	1015	69.9	3444	6	AX145422	AX145422 Sequence
C 7	1015	69.9	4045	1	AF269889	AF269889 Staphyloc
C 8	1015	69.9	4045	6	AX145207	AX145207 Sequence
9	943	64.9	783	6	AX141501	AX141501 Sequence
10	702.5	48.3	250050	1	AL591978	AL591978 Listeria
11	695.5	47.9	5635	6	AX416832	AX416832 Sequence
C 12	686.5	47.2	195269	6	AX417035	AX417035 Sequence
13	686.5	47.2	33050	1	AL596168	AL596168 Listeria
14	686.5	47.2	349980	6	AX417044	AX417044 Sequence
15	657	45.2	894	6	EL1868	EL1868 Mutated DNA
16	656	45.1	894	6	EL1869	EL1869 Mutated DNA
17	654	45.0	894	6	ES4864	ES4864 Process for
C 18	653	44.9	894	6	EL1866	EL1866 Mutated DNA
19	650	44.7	894	6	AR148386	AR148386 Sequence
C 20	650	44.7	894	6	E05337	E05337 DNA encodin
21	650	44.7	894	6	EL1867	EL1867 Mutated DNA
C 22	650	44.7	894	6	EL14762	EL14762 gDNA encodi
23	650	44.7	894	6	EL16095	EL16095 DNA encodin
C 24	650	44.7	894	6	E27521	E27521 Geranyl dip
25	650	44.7	1260	1	BACFDPS	BACFDPS Micrococ
C 26	649	44.7	1080	1	AB003187	AB003187 Mutated DNA
27	645	44.4	894	6	EL1865	EL1865 Streptoco
C 28	644	44.3	894	6	E27520	E27520 Geranyl dip
29	638	43.9	300950	1	AP001516	AP001516 Bacillus
C 30	635	43.7	811	6	AX432600	AX432600 Sequence
31	632.5	43.5	218470	1	BSUB0013	BSUB0013 Bacillus su
C 32	632.5	43.5	282700	1	BACJH842	BACJH842 Bacillus su
33	593	40.8	2268	1	LM0012349	LM0012349 Listeria
C 34	576	39.6	11055	1	AE010637	AE010637 Fusobacte
C 35	556.5	38.3	13536	1	AE010066	AE010066 Streptoco
C 36	555.5	38.2	50463	1	AE014159	AE014159 Streptoco
C 37	551.5	38.0	10165	1	AE006583	AE006583 Streptoco
C 38	551	37.9	876	6	AR112484	AR112484 Sequence
C 39	551	37.9	296750	1	AP003191	AP003191 Clostridi
C 40	549	37.8	10011	6	BD003756	BD003756 Polynucle
C 41	549	37.8	11999	1	AE007421	AE007421 Streptoco
C 42	549	37.8	19059	2	SPNEU1917	SPNEU1917 Streptoco
C 43	547	37.6	10598	1	AE008481	AE008481 Streptoco
44	540.5	37.2	10029	1	AE006320	AE006320 Lactococc
45	535.5	36.9	1548	8	AF425235	AF425235 Abies gra

ALIGNMENTS

RESULT 1

AP004827/c
LOCUS Staphylococcus aureus subsp. aureus MW2 DNA linear BCT 02-JUL-2002
DEFINITION strain:MW2, section 6/10.
ACCESSION AP004827 BA0000033
VERSION AP004827.1 GI:21204509
KEYWORDS Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
SOURCE Staphylococcus aureus subsp. aureus MW2
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A., Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L., Yamamoto,K. and Hiramatsu,K.
TITLE Genome and virulence determinants of high virulence community-acquired MRSA
JOURNAL Lancet 359 (9320), 1819-1827 (2002)
MEDLINE 22040717
PUBMED 12044378
AUTHORS 2 (bases 1 to 333750)
Aoki,K., Oguchi,A., Nagai,Y., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) Akio Oguchi, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/, Tel:81-3-3481-8423, Fax:81-3-3481-8424)
FEATURES
source Location/Qualifiers
1. .333750
/organism="Staphylococcus aureus subsp. aureus MW2"
/strain="MW2"
/db_xref="taxon:196620"
complement(278. .619)
/gene="MW1341"
complement(278. .619)
/gene="MW1341"
/note="ORFID:MW1341"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB95206.1"
/db_xref="GI:21204510"
/translation="MTNITKEVFDNLQESIDVFANKTKLGSSEAKPYLDEYHSKIIDY FKQNDITGNIDFDKLNQPVVPMNFKERYDIYMERKVFHMGYROMKTFKTELKRNA SYQTRLKKNQV"
complement(624. .1283)
/gene="nth"
complement(624. .1283)
/note="ORFID:MW1342"
/codon_start=1
/transl_table=11
/product="endonuclease-like protein"
/protein_id="BAB95207.1"
/db_xref="GI:21204511"
/translation="MVSRRKALEMIDIVIANMPPDAECLEKHNPPFETATVLLSAQCT DYLVRVTEFKKTPEDYLAVSDELMNDIRSIGLYRNKAKNIKLCQSLIDQFN GPOTHTKESLAGVGRKTAIVMVADEPDLAVDTHVERVSKRLGINRWKDNVRO VEDRLCSVIPDRWNRSHQLIFGGRYHCLARKPKDCIPLEDCEGQKRYKASLKE A"
complement(1273. .1959)
/gene="MW1343"
complement(1273. .1959)
/gene="MW1343"
/note="ORFID:MW1343
hypothetical protein, similar to chromosome replication initiation protein dnaH"
/codon_start=1
/transl_table=11
/protein_id="BAB95208.1"
/db_xref="GI:21204512"
/translation="MDKVLKARPVVIRRELLDHYSDGLDQDLVILLIKLIYASETS
NKPSTIELQKQSTMQPRDITMVIVIONLQIRELLELQVQKDEGRFTEYMNLDPPFEKL
SHLKQSQMETREQNSKEFKQLFRVLEDTFARPLSPYEIETLNQWIDVKKHDSALIQ
AALDEANSLNKLSFKYMDRILLNWKNNKVTIDDSRKIREKFNKPRMTHVTVPKPED
WNGENLDCK"
complement(2287. .3579)
/gene="asn"
complement(2287. .3579)
/note="ORFID:MW1344"
/codon_start=1
/transl_table=11
/product="asparaginyl-tRNA synthetase"
/protein_id="BAB95209.1"
/db_xref="GI:21204513"
/translation="MKTITKOAKHLNQDVTIGAWITNKRSSCKIAFLQLRDGTGPMQ
GVVKSVEDEVEFKLAKETOESSLYVTGTITEDNRSDLGYEYQVKSIEVSEAHDP
ITPKNGTEFLMDHRLHWRKQHAVMIRNEVYKATIEFKNGKFTKVDPPHLLTAS
APEGTSELFHTKYFDODAFLSQGLYLEAAMAHGKVFSGPTFRAEKSRYRTPP
FMWIEGMAFTNAESLEIQEYVTHVKSIVLEKLEKLEKLEKLEKLEKLEKLEKLE
ISYDDAIEFKSGEDDIEGDFGAFHETAIANHYDLVPVITNTYTKYKPFYMQPNP
ENEETVLCADLIAPEGYGEIIGGSRVDDLELLEQVRKVEHGLDEEAYSYYLDLRRYGS
VPKCGFGLERTVAWISGVHRETAPFRLNRLYP"
complement(3901. .6594)
/gene="ding"
complement(3901. .6594)
/note="ORFID:MW1345"
/codon_start=1
/transl_table=11
/product="probable ATP-dependent DNA helicase ding"
/protein_id="BAB95210.1"
/db_xref="GI:21204514"
/translation="MGMATYAVVDLETTGNQDDEDDIOIGITFVRNNOIIDYHSMI
RTNLEIPPEIQALTSTEENMQOAPYENQAOEIVDKIDCFVAVHNVDFDLNFIKA
FKDCNIQRPKVVITLFAFKTADSKYSELAEGHITLANHRADEADATATK
LMILAFERFKPLDTLKQLYLQSKQYDLYIFEMVRYQDAKPLDKSYKFFQI
YKQVDFKPTTNGSLKSYKAVDQGLTYRFOQLYLAETILDLQMLHSEKAMIEA
SLGSKSLAYLLAALNIETGKHYMISTNTKLQSKLEKIDPAMNEALNFKINALL
IKSKSDYISGLISQILKDDTSENVNIMKOLLITWITPSGDIQELNKGKQWYF
DQKIEYVPAHDVHYNFIRKNAQNIQIGITNHAHLHSDVENSTYOLFDDCIVDEA
HRLPYALNQVTNELSYADIKYQGLGKIGNEKLLKALDQLEKORILEKLDIAPIDI
FGKASMEHLENEQLSTFIITINDSDVDDIHRPHNVFTFETDLKDLHAIID
KLUNKLEIFNGISHKTSKRLKOLLYLDKFKNIQSKAGHSTISIKNLQSKSTIR
LVYQVAYKDLVTQVLEKSLIFISGTLKFNFHSEFAKQLFNKDHFTEVNTSL
QSAKNTFIPSDVASQYKNDIEVASIVSYIEVTTTSKCLVFTSYKMHMVQ
DMLNELPFEDVYVLTQQQNYKIVQGFNFEDKALLIGTTSFEFSGFQAGIKCV
TAKLPPMKHNAKYLMSSEFTSTFEYVLPDAVTRFROGLRLIRSENDRGIIVSFD
DRLINSYKNFEEQTLNFRQKGDIIQFGKLLRQIQKKKK"
complement(6618. .7589)
/gene="MW1346"
complement(6618. .7589)
/note="ORFID:MW1346"
/codon_start=1
/transl_table=11
/product="truncated hypothetical protein"
/protein_id="BAB95211.1"
/db_xref="GI:21204515"
/translation="MSKYSDQVQLLYLKNKPNYISQSTAESLNSRTAVKVIQDLK
LEGCKIDSVNKHGHLQQLPDWYQIIDQYTKSALDFSEVYDSDISDTQAAKSL
VGNQSSFFLSDQETKGRFRNHWSSSKGGLMWSVLRPNVAFSMISKFNLTALG
IRDAIQHFSDQEVKYPNDIYIDNGKVGCGFLTEIVANNNDGIEATIGICNLQGLE
NFDESIRHRTASVOLHKNKLDYQFLERLQIEIKRYNQFLTLPFSEIREYIAASN
IWNRTLLFTENDQFKGOAIDLQDYSYLIVRDEAGESHRLISADIDF"
complement(7576. .8778)
/gene="MW1347"
complement(7576. .8778)
/note="ORFID:MW1347
hypothetical protein, similar to poly(A) polymerase"
/codon_start=1
/transl_table=11
/protein_id="BAB95212.1"

/db_xref="GI:21204516"
/translation="MDKSLFEQAPLEIQIDQNGFEAYVYGGVDRDVMGRNIHIDI
TTSAPDEISFISHTIPVKEHGTINVFNENYEVTTFRAEEDYDHRPSPGVTFV
RDLYEQSRDFTNNAFAMDTAYKLYIDYFDQDDINNRITRTVGIAERFQEDALRMI
RCLFQSLQSDFIATFEAMRIQYADIKFLSIRIVIELTKLMRGINVEKSPNHLKS
LKAFNYPYIEHLMDQINTEADLELLAIIVSKFDINVLKPLKLSNROVKDINO
YQIMNALPSIIKEQLKMFVYDTHLKNVVAADVIRANDIQGHEPLIVNLOTID
ETLRLPMMHKKMVGNGVLAHLNAGSPWLADVLKLEIALIVTGKVSNEETEILK
WVDNHVKI"

gene complement(8783. .9925)
/gene="MW1348"
CDS complement(8783. .9925)
/gene="MW1348"
/note="ORFID:MW1348
hypothetical protein, similar to lipopolysaccharide
biosynthesis-related pr homolog"
/codon_start=1
/transl_table=11
/protein_id="BAB95213.1"
/db_xref="GI:21204517"

/translation="MKIGITCPSMGSGIIATBELGKLAERGHEVHFTSNIPFRIR
KPLPMFIHOFVENVQYAFQYPPYDITLSTKIAEVIKEYGLDLHMHYAPHAICGIL
AREMSGDIKIMTLHGTDITLVGSDHSLQAGIAFKFEKSDIVTSVSKSLAQETHI
ETNKEIIPYNFVRENEPPTKNTALAKSOFGIAPDEKVLHVSNFROVKRDTIETEF
AKVREKIPSKLILGDPGLVPMRQLKELNVEDVLELGKQCVSEFYQLSDIVLL
SEKESFGLTLEAMKGTVPVIGSNAGGIKVIKRGFTGVVVDGDCDSYAIRLLE
DKVLYNKLQNMADIARFSGELITDQYEYVQKMLNEHNSKGE"
/gene="MW1349"
complement(10167. .10484)
/gene="MW1349"
/note="ORFID:MW1349"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95214.1"
/db_xref="GI:21204518"

/translation="MKSVMOREVDEYIGQFKGYSPPLANLARLLEEVEGELAREIN
HTYGEKKKQSEERATKAEIGDNLFLVLLCLANSMDMTESFNETKEKFNTRDKNRP
ERK"
complement(10820. .11518)
/gene="MW1350"
/gene="MW1350"
/note="ORFID:MW1350"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95215.1"
/db_xref="GI:21204519"

/translation="MSSIFTMIIFYVILMVLPLWAQHKVKNYKEYSOVRSTSGTGR
EVALEILHANGIYDEVVKGEGFTDHYDKPKVYVSLSPANYDRPSVAGTAIAAAHEVG
HAIQHQGYWFLPRAALVPVYANLGSSLSYMIIMLGILITAGISAFGSTALWICAGLM
SLAVLSITLPEVFDASSRAMKOITALINVEKEYKHARKVLSAAAATVVAATAVAV
AELVRIILLARSSD"

Alignment Scores:

Pred. No.:	6,72e-113	Length:	333750
Score:	1448.00	Matches:	286
Percent Similarity:	99.65%	Conservative:	0
Best Local Similarity:	99.65%	Mismatches:	1
Query Match:	99.66%	Indels:	0
DB:	1	Gaps:	0

US-09-925-637-64 (1-287) x AP004827 (1-333750)

Qy 1 MetThrAsnLeuProMetAsnLysLeuLeuValAsnAsnGluLeuSerValala 20
Db 118374 ATGACGAATCACCAGTAATAAATAATAGATGAATGCAATATTCGGTGGC 118315
Qy 21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
Db 118314 ATAAATAATCAGTAATGGATACCTACGCTAGACGAAGAGTAGTGTGTTATTCATTAAATGCT 118255

Qy 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
Db 118254 GGAGCTAAACGCATCCGACCACTTCTGTATTACTCACTTACATTAATACCCGAG 118195
Qy 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
Db 118194 TATGAGTTAGTATGAGAGCGCAATTCACCTAGAAATGATTCAATACATTTAT 118135
Qy 81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
Db 118134 CATGATGACCTACCAGCGATGATATGATGATTCACGAGGAAATTAACAATCAT 118075
Qy 101 LysValTyrGlyGluThrThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
Db 118074 AAAGTATATGCTGAGTGGACTCCGATATTAGCAGGTGATGCTTTTAACTAAAGCATTT 118015
Qy 121 GluLeuIleSerSerAspAspArgLeuThrAspGluValValLysIleLysValLeuGlnArg 140
Db 118014 GAATTTATTCAGGTGATGATAGTAACTGATGAAGTAAATAAATAATTCACACGG 117955
Qy 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
Db 117954 CTGTCATAGCAGCGTGCATGTTGGAATGCTCGCGCTCAATGTTAGATATGCAAGC 117895
Qy 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
Db 117894 GAAGGCCAACCAATTTGAACTTTGAAATGATACACAAAAACACAGGAGCA 117835
Qy 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrLys 200
Db 117834 TTATTAACCTTTGCGGTTATGAGTCAGCAGATATCGCTAATGTCGATGATGCAACTAA 117775
Qy 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspLeu 220
Db 117774 GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCACATTAAGATGATTA 117715
Qy 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsn 240
Db 117714 TTAGACTCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 117655
Qy 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyValGluAspLysLeuThrTyr 260
Db 117654 AATAAAGTACGTACGTGAGTTTATAGGAAAGATGCGCAGAGATGATGATGATGATGAT 117595
Qy 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
Db 117594 CATAGAGCGCAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 117535
Qy 281 LeuLeuGluIleValAspLeu 287
Db 117534 TTATTAGAAATCGTTGATTTA 117514

RESULT 2

AP003134/c
LOCUS AP003134 301550 bp DNA linear BCT 02-JUL-2002
DEFINITION Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
genome, section 6/10.
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS Staphylococcus aureus subsp. aureus N315 (sub_species:aureus N315,
strain:N315) DNA.
SOURCE Staphylococcus aureus subsp. aureus N315
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.

REFERENCE
AUTHORS Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Iian,J., Ito,T., Kanamori,M.,
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
Mikutani-Ui,Y., Takahashi,N., Sawano,T., Inoue,R., Kaito,C.,
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.

Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*
TITLE
JOURNAL Lancet 357 (9264), 1225-1240 (2001)
MEDLINE 21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 301550)
AUTHORS Aoki K., Oguchi A., Hosoyama A., Nagai Y., Kuroda M., Hiramatsu K., Direct Submission
TITLE Submitted (30-JAN-2001) Akio Oguchi, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
JOURNAL (E-mail: oguchi@nitech.go.jp, URL: http://www.bio.nite.go.jp/, Tel: 81-3-3481-8423, Fax: 81-3-3481-8424)
COMMENT On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES
source
 1. 301550
 /organism="Staphylococcus aureus subsp. aureus N315"
 /strain="N315"
 /sub_species="aureus N315"
 /db_xref="taxon:158879"
 complement(249..947)
 /gene="SA1293"
 complement(249..947)
 /gene="SA1293"
 /note="ORFID: SA1293"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB42553.1"
 /db_xref="GI:13701259"
 /translation="MSSIFMTIYIFVLMPLMAQHKVKSNEYKYSQVRSRSGKTGR
 EVALETHANGYDVEVKGEGFLTDHYDKPKVSVLSNPANDRSPVAGTAAGHEVG
 HAQDQGWFLFRFAALVPVANSLSYMIIMLGIIITAGSAPGATAGAGLM
 SLAVFSIVTLVPERDASSRAMKQITAINVNEKEYKHARKVLSRAATYVATAV
 AELVRILLARSD"
 complement(1001..1588)
 /gene="SA1294"
 complement(1001..1588)
 /gene="SA1294"
 /note="ORFID: SA1294"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB42554.1"
 /db_xref="GI:13701260"
 /translation="MTINTFWOYTLVQSWLMLLTCNLIGMIYGIWYGEOLSHTPW
 HKFIVPDSPTAILXIVISILILQKNSIIEALAFVTLFKYGIWAVIMNLIILEQ
 GDITVNLGLVMSHSMVAQAIYFPRFRSMIGISVAMVFLNDIYDFHLPQPYI
 DFITTHVQIGVLSCLSVFGLLLYIELNKLKCK"
 complement(1578..2153)
 /gene="SA1295"
 complement(1578..2153)
 /gene="SA1295"
 /note="ORFID: SA1295"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB42555.1"
 /db_xref="GI:13701261"
 /translation="MSETLNQIKESFIEYLLFQYRFSRIAVVNLNKKVAKLANI
 HFVDTKNHETLETAEGVSHASAIQFTKRNKLMNTNEIFDIYIANHNCAPDQIHEFA
 NVSKREQRLLDVAQLTESPYQTYLHDLNSMAIDRHKHALLIDYLLHNLIDLSQMN
 EKQRFVQLIQLTKLVNKHNFQEDLADD"
 complement(2167..3411)
 /gene="SA1296"
 complement(2167..3411)
 /gene="SA1296"
 /note="ORFID: SA1296"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAB42556.1"

/db_xref="GI:13701262"
 /translation="MEDIYKLIIDDINLQKLENDLSRVNEAITDNDALFILGETLYN
 FGLPQGLEFVRVLYHKYDPDESELLIYFIEGLMSENQTOEALYLSVSPSPKLMLE
 ADLYQOINMVEAIDQLQALELEPNPIIHFALEMLYDGOYLATSEYEVLTETG
 EYQVGNVLFSDMADCSLOSNSYDAIRLYDEINEMTSEDYLLKAYSXDKMLTQCE
 AIKWITLSKDPDYIQGLYIQSLYENKKNYPDAIETKEGRLRSQFTELMTYTCG
 LEIEHDGANEVLLKQALEVNAYOEPLLIISLDYRNEEDYEAILELTYVDEEDLD
 PTFMWHAFAYQOERDKQAQHFELAYPTMTKNTINIDFMSDYFYELIEIQKEQAITYL
 NOLLELEPSNENWHDELSLQS"
 complement(3418..4716)
 /gene="arocA"
 complement(3418..4716)
 /gene="arocA"
 /note="ORFID: SA1297"
 /codon_start=1
 /transl_table=11
 /product="3-phosphoshikimate 1-carboxyvinyltransferase"
 /protein_id="BAB42557.1"
 /db_xref="GI:13701263"
 /translation="MVSEQIIDIISGLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYK
 PLLGDCRRITMDIFRLGLVDIKEDKLVNPGYKAFKTPHOVLVTGNSGTTRLLA
 GLLSGIGIESVLGSDVSGIKRPMDRVLRPLKMDANIEGIEDNYTPIIJKPSVIKGIN
 YOMEVASAQVKSAILFASLESNDTIVIKELDVSRNTEFRHFNPIEAEUSITIT
 PDAIQHIKPADFHVPGDISAFAFVIAALITIPESDVTIHNVGINPRSGIIDVERKM
 GNIQFNQTFGAAPTASIRIQYTPMLQPIITIEGELPKAIDELPVALICLTQAVGTST
 IKDAELKVKETNRIDTTADMLNLGLFELQPTNDGLIIHPSEFKTNATVDSLTDHRRIG
 MMLAVASLSSSEPVKIQFDVAVNSFPGLPKLKLENEG"
 complement(4726..5790)
 /gene="arob"
 complement(4726..5790)
 /gene="arob"
 /note="ORFID: SA1298"
 /codon_start=1
 /transl_table=11
 /product="3-dehydroquinate synthase"
 /protein_id="BAB42558.1"
 /db_xref="GI:13701264"
 /translation="MKLQTYTPSNNYPIFVEHGAIDHISTYIDQFQSPILIDEHVNO
 YFADKFDIDILSYENVHKVITIPAGEKTKTEQOYQETLEYILSHHVTRNTAIIVGGGAT
 GPFAGVAAITLLRGVHFIOVPTTILAHDSVGVKGINSKQKNLIGAFYRTPAVIYD
 LPLKTLPEQLILSGVAEYVYKHALNGESTQIEIQHFQKDELQSLQSMKDYIANGI
 ETKLDIVADEKEQGVKFLNCHTFGHAVEYNHKAHGHAVHIGIYOFIVANILFN
 SNHDIOHYNILTKGLPLETITDIDFETIYQYMLSDKKNDKQGVOMVLKHPGDIV
 QHIDQTTQHAECQLKTYFK"
 complement(5816..6982)
 /gene="aroc"
 complement(5816..6982)
 /gene="aroc"
 /note="ORFID: SA1299"
 /codon_start=1
 /transl_table=11
 /product="chorismate synthase"
 /protein_id="BAB42559.1"
 /db_xref="GI:13701265"
 /translation="MRLTSGSHGHPQLTVIVEGIPANLEIKVEDINKEMFKRQGGY
 RGRMQIEKDTVEIVSGVNGYTLGSPITMVVYNDFTWHRKIMGAAPSEERENMK
 RTTIPRPGHADLVGMKYNHRDLNRNLSRSARETAARVAGALCKVLQQLDIDY
 SRVTEIGKDKDFDYSKTFKANLDRNVRVDDSTAQAMRDKIDEAKNKGSIIGVY
 QVVENMPVGVSVYHYDKLQAGVVSINAKGVSGFGFKAAERKPSGEIQDEI
 LNSEIGYRGSNHNHGLGEGMSNGMPLVNGVMKPIPTLYAPLNSVDINTDEKAT
 IERSDSCAVPASIVCEHVVAFEIARAKALLEEFOQSNHIEQLQOOIADRRQLNVEF"
 complement(7368..7568)
 /gene="SA1300"
 complement(7368..7568)
 /gene="SA1300"
 /note="ORFID: SA1300"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAB42560.1"
 /db_xref="GI:13701266"
 /translation="MKNIIISITILGIMFLKMLLYGAIFLDKPLNPTIKIIFILTI
 YIFYVLVKEIIFLKSYNKSA"

Db	63770	GAACCTTATTTC	CAAGTGATGATAGATTAACATGATGAAGTAAAAATAAAAGTTCTTACAACGG	63711
QY	141	LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer	160	
Db	63710	CTGTC	AAATGACCAATTCCTTGA	63651
QY	161	GluglyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla	180	
Db	63650	GAAGGCCAACCAATTCCTTGA	63591	
QY	181	LeuLeuThrPheAlaValMetSerIleAlaAspIleAlaAsnValAspAspThrLys	200	
Db	63590	TTATTTAAC	63531	
QY	201	GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu	220	
Db	63530	GAACATTTAGAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAGATGATTTA	63471	
QY	221	LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn	240	
Db	63470	TTAGACTGCTATGGTGATGAAGCAAAAGTTAGTAAAAAAGTGGCGAGCATCTTGA	63411	
QY	241	AsnLysSerThrTyrValSerLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr	260	
Db	63410	ATAAAAAGTACGTACGTGAGTTTATTAGGAAAGATGGCGAGAGATATAATGAC	63351	
QY	261	HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis	280	
Db	63350	CATAGAGCGAGCAGTGGATGAACCTAGCAAAATGATGAACAATTCATACAAAACAC	63291	
QY	281	LeuLeuGluIleValAspLeu	287	
Db	63290	TTATTAGAATCGTTGATTTA	63270	
RESULT 3	AP003362/c	LOCUS	DEFINITION	AP003362 346900 bp DNA linear BCT 07-FEB-2002
ACCESSION	AP003362	VERSION	BA000017	Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete sequence, section 5/9.
KEYWORDS	AP003362.2	SOURCE	GI:14247083	
ORGANISM	Staphylococcus aureus subsp. aureus Mu50 (sub-species: aureus Mu50, strain: Mu50) DNA.	REFERENCE	1	Staphylococcus aureus subsp. aureus Mu50 (sub-species: aureus Mu50, strain: Mu50) DNA.
AUTHORS	Cui, R., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, R., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kihara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.	TITLE	Whole genome sequencing of methicillin-resistant Staphylococcus aureus	
JOURNAL	Lancet 357 (9264), 1225-1240 (2001)	MEDLINE	21311952	
REFERENCE	2 (bases 1 to 346900)	AUTHORS	Ohta, T.	
TITLE	Direct Submission	JOURNAL	Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College of Medical Technology and Nursing, Department of Medical Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan (E-mail: tohtadesakura.cc.tsukuba.ac.jp. Tel: 81-298-53-3454, Fax: 81-298-53-3454)	
COMMENT	On May 29, 2001 this sequence version replaced gi:13875626.	FEATURES	Location/Qualifiers	
source	1. .346900			
	/organism="Staphylococcus aureus subsp. aureus Mu50"			
	/strain="Mu50"			
	/sub_species="aureus Mu50"			
	/db_xref="taxon:158878"			

[illegible]

```
Db 2193 GGTGGTAAAGAAATCAGACAGCATATATTATTAAACACTTAAATAATGCTTACAAAGAT 2134
QY 61 TyrGluLeuGluMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
Db 2133 TATCAACAGGACTAAATAGTGTGTTAGCATTTGGAATGATTCATATCTTCTTTAAAT 2074
QY 81 HisAspAspLeuProAlaMetAspAspTyrArgGlyLysLeuThrAsnHis 100
Db 2073 CATGATGATTTACCAGCAATGGAATAGCATTTACCGTAGAGAAATTAACAAATCAT 2014
QY 101 LysValTyrGlyClnTrpThrAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 120
Db 2013 AAAGTTATGTTGTAATGGAAGCCATCTGCTGGTGAATGATTTATTAACAAAGCTTTT 1954
QY 121 GluLeuIleSerSerAspArgLeuThrAspGluValLysLysValLeuGlnArg 140
Db 1953 GAATTTAGTGTGTAATGGAAGCCATCTGCTGGTGAATGATTTATTAACAAAGCTTTT 1954
QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuGluMetIleHisThrTyrSerLeuIle 160
Db 1953 GAATTTAGTGTGTAATGGAAGCCATCTGCTGGTGAATGATTTATTAACAAAGCTTTT 1954
QY 161 GluGlyClnProIleAspLeuThrAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 180
Db 1833 GAAGGGAAGTCAATTCGTTTAGAATCAATTCAGTAACTAAGACGCGCT 1774
QY 181 LeuLeuThrPheAlaValMetSerAlaIleAlaAspIleAlaAsnValAspAspThrThrLys 200
Db 1773 TTACTAAATTTTTCAGTTATGCTGCGGTAGACATTTGCTCAAGTAAAGCAAAATATTGCT 1714
QY 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheClnIleLysAspLeu 220
Db 1713 AAGAATTTAGTGTGTAATGGAAGCCATCTGCTGGTGAATGATTTATTAACAAAGCTTTT 1954
QY 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsn 240
Db 1653 CTGGATGTGTGTTGATGAATCAAACTTGGCAAAAGTAGTAGGCAAGTATAGTAAAT 1594
QY 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
Db 1593 CATAAAGTACATATGCTTTTCTTACTTGGANAAGAGGAGCAGAGAGAAAGTAAACAT 1534
QY 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluClnPheAsnThrLysHis 280
Db 1533 CATCAATATCTGCTATGATGATCAAAATTCGTCAATCAATATGATCTCTGAA 1474
QY 281 LeuLeuGluIleValAspLeu 287
Db 1473 TTAAGTGTATTTGTAGATTTA 1453
```

```
RESULT 8
AXI45207/c 4045 bp DNA linear PAT 31-MAY-2001
LOCUS Sequence 3929 from Patent WO0134809.
DEFINITION AXI45207
ACCESSION AXI45207.1 GI:14283772
VERSION 1 (bases 1 to 4045)
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 4045)
AUTHORS Kimmerly, W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 3929 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
    source
        1. .4045
        /organism="synthetic construct"
        /db_xref="taxon:32630"
        /note="synthetic nucleic acid sequence"
BASE COUNT 1249 a 691 c 536 g 1569 t
ORIGIN
```

```
Alignment Scores: 5.63e-78 Length: 4045
Pred. No.: 1015.00 Matches: 197
Score: 82.93% Conservative: 41
Percent Similarity: 82.93% Mismatches: 49
Best Local Similarity: 68.64% Indels: 0
Query Match: 69.86% Gaps: 0
DB: 6
US-09-925-637-64 (1-287) x AXI45207 (1-4045)
QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
Db 2313 ATGAAGAACTACAGATGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2254
QY 21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
Db 2253 ATACAATCATCCATCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 2194
QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 2193 GGTGGTAAAGAAATCAGACGAGCATATTTATTAACACTTAAATGCTTAAACAAGAT 2134
QY 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
Db 2133 TATCAACAGGACTAAATAGTGTGTTAGCATTTGGAATGATTCATATCTTCTTTAAAT 2074
QY 81 HisAspAspLeuProAlaMetAspAspTyrArgGlyLysLeuThrAsnHis 100
Db 2073 CATGATGATTTACCAGCAATGGAATAGCATTTACCGTAGAGAAATTAACAAATCAT 2014
QY 101 LysValTyrGlyClnTrpThrAlaIleAlaLeuAlaGlyAspAlaLeuLeuLeuThrLysAlaPhe 120
Db 2013 AAAGTTATGTTGTAATGGAAGCCATCTGCTGGTGAATGATTTATTAACAAAGCTTTT 1954
QY 121 GluLeuIleSerSerAspArgLeuThrAspGluValLysLysValLeuGlnArg 140
Db 1953 GAATTTAGTGTGTAATGGAAGCCATCTGCTGGTGAATGATTTATTAACAAAGCTTTT 1954
QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuGluMetIleHisThrTyrSerLeuIle 160
Db 1893 CTTTCAAAAGCAAGTGGACATTTGGANAAGAGGAGCAGAGAGAAAGTAAACAT 1834
QY 161 GluGlyClnProIleAspLeuThrAlaIleAlaLeuGluMetIleHisThrLysThrLys 180
Db 1833 GAAGGGAAGTCAATTCGTTTAGAATCAATTCAGTAACTAAGACGCGCT 1774
QY 181 LeuLeuThrPheAlaValMetSerAlaIleAlaAspIleAlaAsnValAspAspThrThrLys 200
Db 1773 TTACTAAATTTTTCAGTTATGCTGCGGTAGACATTTGCTCAAGTAAAGCAAAATATTGCT 1714
QY 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheClnIleLysAspLeu 220
Db 1713 AAGAATTTAGTGTGTAATGGAAGCCATCTGCTGGTGAATGATTTATTAACAAAGCTTTT 1954
QY 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsn 240
Db 1653 CTGGATGTGTGTTGATGAATCAAACTTGGCAAAAGTAGTAGGCAAGTATAGTAAAT 1594
QY 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
Db 1593 CATAAAGTACATATGCTTTTCTTACTTGGANAAGAGGAGCAGAGAGAAAGTAAACAT 1534
QY 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluClnPheAsnThrLysHis 280
Db 1533 CATCAATATCTGCTATGATGATCAAAATTCGTCAATCAATATGATCTCTGAA 1474
QY 281 LeuLeuGluIleValAspLeu 287
Db 1473 TTAAGTGTATTTGTAGATTTA 1453
RESULT 9
AXI41501
```

Db 601 GTAGGCAGTCATATACTAATAACAAAGAAGTACTATTGTTCTTTACTTGCAAAAAGGA 660

Qy 254 AlaGluAspLysLeuThrTyHisArgAspAlaAlaValaspGluLeuThrGlnIleAsp 273
|||||:||||||| :||||| :||||| :|||||

Db 661 GCAGAGAAAGATTAAACAATCATCAATTCATCTGCTGTGAACGTCTTAATCAAATTTCT 720
::|||:||||| ||| :|||||

Qy 274 GluInPheAsnThrLysHisLeuGluGluIleValaspLeu 287
::|||:||||| ||| :|||||

Db 721 GATCAATATGACTTCTCGAATTAAGTAGATATTGTAGATTTA 762

RESULT 10
AL591978 250050 bp DNA linear BCT 06-JUN-2002
LOCUS Listeria monocytogenes strain EGD, complete genome, segment 6/12.
DEFINITION AL591978 AL591824
ACCESSION AL591978.1 GI:16410540
VERSION
KEYWORDS
SOURCE Listeria monocytogenes.
ORGANISM Listeria monocytogenes
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

REFERENCE 1
AUTHORS Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A.,
Baquero, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T.,
Charbit, A., Chetoui, F., Couve, E., de Daruvar, A., Dehoux, P.,
Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L.,
Dussurget, O., Entian, K.D., Psihi, H., Portillo, F.G., Garrido, P.,
Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J.,
Jackson, D., Jones, L.M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F.,
Kurapkat, G., Madueno, E., Maitournam, A., Vincent, J.M., Ng, E.,
Nedjari, H., Nordstrek, G., Novella, S., de Pablo, B., Perez-Diaz, J.C.,
Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N.,
Tierrez, A., Vazquez-Boland, J.A., Voss, H., Wehlend, J. and Cossart, P.
Comparative genomics of *Listeria species*
Science 294 (5543), 849-852 (2001)
21537279

TITLE Listeria monocytogenes
JOURNAL MEDLINE
PUBMED 11679669

REFERENCE 2 (bases 1 to 250050)
AUTHORS Glaser, P., Frangeul, L. and Rusniok, C.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cedex 15, FRANCE

COMMENT E-mail: pglaser@pasteur.fr
Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.

FEATURES
source
1..250050
 /organism="Listeria monocytogenes"
 /strain="EGD-e"
 /db_xref="taxon:1639"
210..753
 /gene="lmo1125"
210..215
 /gene="lmo1125"
220..753
 /gene="lmo1125"
 /codon_start=1
 /protein_id="CAC99203.1"
 /db_xref="GI:16410541"
 /db_xref="SPTRMBL:O8X7Z4"
 /transl_table=11
 /translation="MLIPSLKSTAKELSIINDFNVEIVTIINKEETFEILYCGTLKEI
EYDQLITGAYEPILPKIIAKLSIGKEILLYDGAVGYDSMFCDFEFEBATQNRLKIK
YPNNLSNTVRIISIGIDYESEKEEYFENGHVLLIDDRHPWEQVKTDGDFDLEIFT
ADWNGDSLIIITEELA"
join(764..769,778..1230)
 /gene="lmo1126"
764..769
 /gene="lmo1126"
778..1230
 /gene="lmo1126"
 /note="similar to E. coli YjaB protein"

gene
RBS
CDS

gene
RBS
CDS

```
/protein_id="CAC99204.1"
/db_xref="SPTREMBL:Q8Y723"
/translation="MILINKRAENQYQMLATWEKSVIATHDFLAAEDRQFYKEQIP
LELDNNELFLAFAGEETVFGSGTSERELDMFLNPIATNGYGSQILWLIENKRLN
VDNEQENATRYLKHGFFVSSRDLDFGPNPILHLQLKERVNGI"
1208..1213
/gene="lmo1127"
RBS
1208..1213
/gene="lmo1127"
1220..1692
/gene="lmo1127"
1220..1675
/gene="lmo1127"
/codon_start=1
/translation="MEYEQQLQNELGELELIQBELONKEYRGMIVKAKDGTITRRL
AKTDRKAGYFAFWEKDFGKNCPCPDENDCADFLCIVVIGDGLNGLFVPPROCLIDK
GILTSEASIGKMAAREVPTWCQDLNKTAKTKTQAWQYVFNYSIRKEDH"
1660..1692
/gene="lmo1127"
terminator
complement(1689..2711)
/gene="lmo1128"
complement(1689..2711)
/gene="lmo1128"
/codon_start=1
/translation="MAKNGVDITKMKPKITIIISAIVILVAILFIGNVNMETRVA
IPTGGLSAVTTPKHGPKGIIIVFHGDAEATONGYKPLMERFAKQGVISVM
DKLGVKSSGNLWLNQSDDRANVYQVIEHMKVYKYPDSYAKIGLWASQAGWVPPRAM
NANNEIDFSLAAPALNMRQGEYTCQAKDQKGTKEVKQQAQNFDTSDALISRKN
TYSYKONGKEEMTPGRYVFIKNNQDADLDSKVTKIYLALAEKDNKVDSAETK
DVTQIQIKENLETQIPNVSHQINPAINSELTNIVGLMVPYFLVDNAYLDICE
QVYSKQ"
complement(2666..3280)
/gene="lmo1129"
complement(2666..3280)
/gene="lmo1129"
/notes="similar to unknown proteins"
/codon_start=1
/translation="MSEWTIIGWLLAFKVRKGMILAILGLIISNLVIMISLIGTY
GIATVILQPLNPKYKLVFVWAGVYVGLARYLEQVLDHDIARLAIIRERFATLR
KIGPARLSKSKSGDLVAAITVDVEALEVFPAHTISPVFALGTITATVGFGLSYDVLG
AIIILVQIMGVVLPWIPISYKRNKIKGTATQQAVGLNQMNMENIASLQDIFQFKIGE
ERLKLQTEKLNQYKRIKQSGELQILGEWLIGTATIIILVGLSGLQPLETVLI
GTVLSLSFGSVLALNALTALITTFASGRLYALFEKPVVFLNGLQELSDPESAEI
DKVFSHGDQPLNLNLSLALPKGVWLGIGGSGSGKSTLVKLLMRYWDPDGVKLN
NELPKITESLHRELEGVMEQRFLEFDTIGNIRLCKKDATLEEVKEAKKAIDTWI
EVLPEGYETIIGGSRNLSDGERQIGRLARLFLHDAPLLLDDEPTSNLDTINQATIN
TURSEQDKTVLVISHRATVLDIAEQFLFIENGALKRAAK"
complement(Join(7604..7637,7645..7953,7961..7966))
/gene="lmo1133"
terminator
complement(7604..7637)
/gene="lmo1133"
7615..7637
/gene="lmo1132"
7615..7637
/gene="lmo1132"
complement(7645..7953)
/gene="lmo1133"
/notes="similar to B. subtilis Yjcs protein"
/codon_start=1
/translation="MYLKENTVLYCTAVIQTQTKVSYEQLVHEHLALRAKTVKEPCG
```

Alignment Scores:
Pred. No.: 7,64e-49
Score: 702.50
Percent Similarity: 69.29%
Best Local Similarity: 52.86%
Query Match: 48.35%

Length: 250050
Matches: 148
Conservative: 46
Mismatches: 81
Indels: 5

```
complement(4004..4009)
/gene="lmo1130"
Join(4233..4237,4245..5960)
/gene="lmo1131"
4233..4237
/gene="lmo1131"
4245..5960
/gene="lmo1131"
/notes="similar to ABC transporters, ATP-binding proteins"
/codon_start=1
/translation="MIDKRLFOLVEKKSLVLLILFRVLSGLMGLWVFAQQLTHYL
EGONVDMWLGIVLVLYGKAILTKIAEKIYQASNELRSLMRRAVMEKAFRLGNE
GQLPASTLTQLAVDGEIEQIYYSRFELPQLFYCLIASLMIFGSLVGFAMQPAIVLLIC
MPMIPIVIMAMKIAKRLISGYSWSDYTNLGRFHEHLSGLSILKAQQDKYKQDEIVS
DAEFFIPMQLGSLFVAMNGISACISKLFAYLEKQVYGSVELAKPLEKIVHHILTY
TYEGGAIRKQAVLVNDHGYLYXANSIRENLIGNPANDSDLANVLEOVSLADFVKL
ASLSGPAIRKQAVLVNDHGYLYXANSIRENLIGNPANDSDLANVLEOVSLADFVKL
PEQNLNLEDGNSLGGORQLLALRLQAEVYVFEITSGVDESKEKILLVILQ
ELAKEKIVLFIHRLYNVLGADQVDFDAGKLLLEVASPERLQOESNYFNKYLEEAL
LKGGK"
5948..5953
/gene="lmo1132"
5948..5953
/gene="lmo1132"
5962..7608
/gene="lmo1132"
5962..7608
/gene="lmo1132"
/notes="similar to ABC transporters, ATP-binding proteins"
/codon_start=1
/translation="CAC99210.1"
/db_xref="GI:16410548"
/db_xref="SPTREMBL:Q8Y7Y7"
/translation="MSEWTIIGWLLAFKVRKGMILAILGLIISNLVIMISLIGTY
GIATVILQPLNPKYKLVFVWAGVYVGLARYLEQVLDHDIARLAIIRERFATLR
KIGPARLSKSKSGDLVAAITVDVEALEVFPAHTISPVFALGTITATVGFGLSYDVLG
AIIILVQIMGVVLPWIPISYKRNKIKGTATQQAVGLNQMNMENIASLQDIFQFKIGE
ERLKLQTEKLNQYKRIKQSGELQILGEWLIGTATIIILVGLSGLQPLETVLI
GTVLSLSFGSVLALNALTALITTFASGRLYALFEKPVVFLNGLQELSDPESAEI
DKVFSHGDQPLNLNLSLALPKGVWLGIGGSGSGKSTLVKLLMRYWDPDGVKLN
NELPKITESLHRELEGVMEQRFLEFDTIGNIRLCKKDATLEEVKEAKKAIDTWI
EVLPEGYETIIGGSRNLSDGERQIGRLARLFLHDAPLLLDDEPTSNLDTINQATIN
TURSEQDKTVLVISHRATVLDIAEQFLFIENGALKRAAK"
complement(Join(7604..7637,7645..7953,7961..7966))
/gene="lmo1133"
terminator
complement(7604..7637)
/gene="lmo1133"
7615..7637
/gene="lmo1132"
7615..7637
/gene="lmo1132"
complement(7645..7953)
/gene="lmo1133"
/notes="similar to B. subtilis Yjcs protein"
/codon_start=1
/translation="CAC99211.1"
/db_xref="GI:16410549"
/db_xref="SPTREMBL:Q8Y7Y6"
/translation="MYLKENTVLYCTAVIQTQTKVSYEQLVHEHLALRAKTVKEPCG
```


DB:	1	Gaps:	2
US-09-925-637-64 (1-287) x AL591978 (1-250050)			
Qy	8	LysLeuileaspGluValasnAsnGluSerValalalleAsnLysSerValMetasp	27
Dd	225970	AAGTGAATGATGAG-----TCGCTTTTAAAGAAATAACGAGCGAATAATCGAA	226020
Qy	28	ThrGlnLeuGluGluSerMetLeuTySerLeuAsnAlaGlyGlyLysArgIleArgPro	47
Dd	226021	CCTAGACTAAAAGAGTCCATTATATTCATTCAACGACGCGGAAACGGATTCGTCCA	226080
Qy	48	ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyGluLeuGlyMetLysSer	67
Dd	226081	ATGCTAGTTTTTCGTCACGCTTCAAGCTTTAAAGTAATCCGCCTTTAGTGTAATAAAC	226140
Qy	68	AlatlealeuGluMetIleHisThrTySerLeuileHisAspAspLeuproAlaMet	87
Dd	226141	GCAACACGGTTAGAANAATGATTCAACATACAGCCTTATTCAATGATGATCTACCAGCAATG	226200
Qy	88	AspAsnAspAspTyArgArgGlyLysLeuThrAsnHisLysValTyGlyGluTriphr	107
Dd	226201	GATPAACCATGATATATCTCGAGGCAATGGACTAATCATAAAGTTTTCGGCATGCCAACT	226260
Qy	108	AlatleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuileSerSerAsp	127
Dd	226261	CGCATTTTGGCAGGAGATGCTTACTAACGCTCGCTTTCTATTCTTAGCTGAAGACGAT	226320
Qy	128	ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis	147
Dd	226321	AAITTAFTCTTTTGAGACACCATGCTTGTGATTAAACCAANATAGTTTAGTAGGGGTGCA	226380
Qy	148	ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleaspLeu	167
Dd	226381	GAAGGATGCTTGGTGTCAACTTCAGACTTGGAGCGGAAAAACAACAGTACGCGTA	226440
Qy	168	GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet	187
Dd	226441	GAAGAGTTATCATCTCAATTCATGTCACGAAAAACGGGTGAATTAATTAATTTATGCTGAACC	226500
Qy	188	SerAlaalaaspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer	207
Dd	226501	TCGACACAAAATTCGGAAGCTGATCCAGAACAAACGACTTACGAAATTTTTCGA	226560
Qy	208	TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCystYrGlyAspGlu	227
Dd	226561	GAGAAATTCGGATGGATTTCAAATTAGCCAGCATATTTAGATGTAATTTGCTGATGAA	226620
Qy	228	AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer	247
Dd	226621	ACGAAATTCGGTAAAAAGACAGCGCCGCGCTTTCTGTAATAAAGTACCTATCCCCGA	226680
Qy	248	LeuLeuLysAspGlyAlaGluAspLysLeuThrThrHisArgAspAlaAlaValasp	267
Dd	226681	TTACTCACGTTGATGGGCAAAAGGCGCATTAATGAGCATGTTACGATTTGCAAGTCA	226740
Qy	268	GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValaspLeu	287
Dd	226741	CGCGTTTCAGGGCATGAT----TTCCGATCATGAATTCCTCTGAAACTTCTGCTGATTTA	226794
RESULT 11			
AX416832		LOCUS	5635 bp DNA linear PAT 14-JUN-2002
DEFINITION		Sequence 3823 from Patent WO0228891.	
ACCESSION		AX416832	
VERSION		AX416832.1 GI:21449341	
KEYWORDS			
SOURCE		Listeria monocytogenes ATCC 19115.	
ORGANISM		Listeria monocytogenes ATCC 19115	
REFERENCE		Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.	
AUTHORS		Glaser,P. and Kunst,F.	
TITLE		Listeria innocua, genome and applications	

JOURNAL Patent: WO 0228891-A 3823 11-APR-2002;
 FEATURES Pasteur Institut (FR)
 source Location/Qualifiers
 1. 5635
 /organism="Listeria monocytogenes ATCC 19115"
 /db_xref="taxon:176281"
 BASE COUNT 1896 a 927 c 1238 g 1574 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.8e-50 Length: 5635
 Score: 695.50 Matches: 148
 Percent Similarity: 68.93% Mismatches: 45
 Best Local Similarity: 52.86% Mismatches: 82
 Query Match: 47.87% Indels: 5
 DB: 6 Gaps: 2

US-09-925-637-64 (1-287) x AX416832 (1-5635)

Qy	8	LysLeuIleAspGluValAsnAsnGluLeuSerValalaIleAsnLysSerValMetAsp	27
Db	2914	AAAGTGATTGATGAG-----TCGCTTTTAAAGAATAAGCAGGCGAAATATCGGA	2964
Qy	28	ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro	47
Db	2965	CCTAGACTAAAGAGTCCATGTTATATCTGTTCAAGCAGGTGGAACGGATTCGTCCC	3024
Qy	48	ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer	67
Db	3025	ATCCTAGTTTTTCTACACTCAAGCTTAAAGTAATCCCGCTTTTATAGTGTGAAACA	3084
Qy	68	AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet	87
Db	3085	GCAACGCGTTAGAAATGATTCAATAGCTGATCCTGATCCATGATGATTACCAGCAATG	3144
Qy	88	AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr	107
Db	3145	GATAATGATGACTATCGTCGCGCAATATACTAATCATATAAGTTTTTGGGATGCAACT	3204
Qy	108	AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp	127
Db	3205	GCATTTTGGCAGGAGACGCTTTACTAACGCTTCTTTCTATTTTACGGAAGACGAA	3264
Qy	128	ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis	147
Db	3265	AAATTATCTTTTGAACACCGCATAGCTTGTGATTAATCAAAATAGTTTTAGTAGTGCGCA	3324
Qy	148	ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu	167
Db	3325	GAAGGAATGTTGGCGGTCAACTTGCAGATATGAGGCGAGAAAAACAACAGTGACATTA	3384
Qy	168	GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet	187
Db	3385	GAAGAACTATCATCCATCCATGACCGGAAACAGCGGAATTAATATTTTTCGCGTAAC	3444
Qy	188	SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer	207
Db	3445	TCGTGTCGGAATTTGCAGAGAGCTGACCCGGAACAACGAAACATACGGATTTTGGG	3504
Qy	208	TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuAspCystyrGlyAspGlu	227
Db	3505	GAATAATATGGGATTGGATTCCCAATCAGCGACGATATTTTACATGCTTATTTGGTGATGA	3564
Qy	228	AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnLysSerThrTyrValSer	247
Db	3565	ACAAAATGGGTAAAAGACAGCGGTCGACGCTTTCTGAACAAAAAGTACCTATCCCGA	3624
Qy	248	LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp	267
Db	3625	TTACTCAGCTTGATGGTGCACAAAGGCGCATTAATGACGATGTTCCGATTCGAAAGTCA	3684
Qy	268	GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu	287


```
Db 193288 TCACCGCGAAATCGCTGAAGCAACTCCAGAACAAACCAACGATTACGAATTTTGCA 193347
Qy 208 TyrHisLeuGlyMetMetPheGlnIleLysAspLeuLeuAspCysTyrGlyAspGlu 227
Db 193348 GAAATATCGCATTTGGTTTTCAAATAGCGACGATATTTAGATTAATTTGGCGATGA 193407
Qy 228 AlalysLeuGlyLysLysValGlySerAspLeuGluAsnLysSerThrTyrValSer 247
Db 193408 ACAAAATGGGTAAAGACAGAGGGGTGACGACCTTTCTGTAATAAAGTACCTATCCCGGA 193467
Qy 248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
Db 193468 TTACTCACATCTCGAAGGTGCTTAACGGGCATTAATAGCAATGTTTCAATTCGAAATGA 193527
Qy 268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
Db 193528 GCGCTTTCAGGACATGAT-----TTCGACGATGAAATTCCTTAAACTTGCTGATTTA 193581
```

RESULT 15

```
1868
CUS
DEFINITION Mutated DNA encoding farnesyl-2-phosphate synthase.
ACCESSION E11868
VERSION E11868.1 GI:22025489
KEYWORDS JP 1996214877-A/4.
SOURCE Bacillus stearothermophilus.
ORGANISM Bacillus stearothermophilus.
REFERENCE 1 (bases 1 to 894)
AUTHORS Kolke,A., Obata,A., Nishino,T., Onuma,S., Nakazawa,T., Ogura,K. and Furiyama,T.
TITLE MUTANT FARNESYL DIPHOSPHATE SYNTHETASE CAPABLE OF SYNTHESIZING GERANYLGERANYL DIPHOSPHATE AND DNA CODING THE ENZYME
JOURNAL Patent: JP 1996214877-A 4 27-AUG-1996;
COMMENT TOYOTA MOTOR CORP
OS Bacillus stearothermophilus
PN JP 1996214877-A/4
PD 27-AUG-1996
PF 14-FEB-1995 JP 1995025253
PI KOIKE AYUMI, OBATA ATSUO, NISHINO TOKUZO, ONUMA SHINICHI, PI
NAKAZAWA TAKESHI, OGURA KYOZO, FURUYAMA TANETOSHI PC
C12N9/12,C07H21/04,C12N1/21,C12N15/09,C12P11/00/G01N33/50, PC
(C12N9/12),
PC C12R1:19,(C12N1/21,C12R1:19),(C12N15/09,C12R1:07): CC
strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FH source 1..894
FH FT source 1..894
FT Location/Qualifiers
FT source 1..894
/db_xref="taxon:1422"
/organism="Bacillus stearothermophilus"
```

FEATURES

```
source 1..894
BASE COUNT 207 a 243 c 274 g 170 t
ORIGIN
```

Alignment Scores:

```
Pred. No.: 5.95e-48 Length: 894
Score: 657.00 Matches: 133
Percent Similarity: 68.10% Conservative: 57
Best Local Similarity: 47.67% Mismatches: 83
Query Match: 45.22% Indels: 6
DB: 6 Gaps: 2
```

US-09-925-637-64 (1-287) x E11868 (1-894)

```
Qy 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValala 20
Db 1 ATGGCGGACGCTTTCAGTTGAACAGTTCTCAACGAGCAAAACAGGCGGTGGAACACGC 60
Qy 21 IleAsnLysSerValMet-----AspThrGlnLeuGluSerMetLeuTyr 36
```

```
Db 61 CTCCTCCCGCTTATAGACGCTTAGAAGGGCGGCGGAAGCTGAAAAAGCGGATCGCGTAC 120
Qy 37 SerLeuAsnAlaGlyGlyArgIleArgProValLeuLeuLeuLeuLeuLeuAspSer 56
Db 121 TCATTTGGAGCGCGGCGGCAACGAATCCGTCGTTGCTGCTCTCTCCACCGCTTCGGCGG 180
Qy 57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
Db 181 CTCGCAAAAGACCCCGCGCTCGGATTGCTCGCTCGCGGATTGAAATGATCATCAGC 240
Qy 77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgLys 96
Db 241 CACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGCGCAG 300
Qy 97 LeuThrAsnHisLysValTyrGlyThrAlaIleLeuAlaGlyAspAlaLeuLeu 116
Db 301 CCGACGACCATATAAGTGTTCGCGGAGCGGATGGCCATCTTGGCGGGGACGGGTGTG 360
Qy 117 ThrLysAlaPheGluLeuIleSer-----SerAspAspArgLeuThrAspGluValLys 134
Db 361 ACGTACGCGTTCATTAATTCATCACCAGAAATCGACGATGAGCGCATCCCTCTCCGCGG 420
Qy 135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGln 154
Db 421 CTTGGGCTCATCGAACGCGTGGCGAAAGCGCGGTCGCGAAGGAGGATGGTCGCGGTCAG 480
Qy 155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174
Db 481 CGACCGGATATGGAAGGAGAGGGGAAACGCTGACGCTTCGGACCTCGAATACATCAT 540
Qy 175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194
Db 541 CGGCATAAAACCGGAAATGCTGCAATACAGCGTGCACGCGCGCCTTGATCGCGCGC 600
Qy 195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214
Db 601 GCTGATGCCCGCAACCGCGGAGGCTTGACGAATTCGCGCGCCCATCTAGCGCTTCCTTT 660
Qy 215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLeuLysLysVal 234
Db 661 CAATTCGCGATGATATTCGATATTGAAGGGGCGAGAAAGAAATCGCAAGCGGTC 720
Qy 235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
Db 721 GGCAGCGACCAAGCAACAAACAGCAGTATCCAGCGTTGCTGCTGCGCGCGCG 780
Qy 255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
Db 781 AAGGAAAAGTTGACGTTCCATATCGAGGCGCGCAGCCCATTTACGGAACGCCGAC 837
```

Search completed: May 30, 2003, 21:30:48

Job time : 3210 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 19:09:11 ; Search time 276 Seconds
(without alignments)
2341.752 Million cell updates/sec

Title: US-09-925-637-64

Perfect score: 1453

Sequence: 1 MTNLPNMKLLDEVNLSVA.....ELTQIDBQFNKHLLEIVDL 287

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Deiop 6.0 , Delext 7.0

arched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US09525637/runat_23052003.174817.23416/app_query.fasta_1.455
-DB=N_Geneseq_101002 -QWTF=fastap -SUFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdd
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09525637 -CGN_1.1_263 -runat_23052003.174817.23416 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1453	100.0	861	22	AAS00821	S. aureus HGS072 e
2	1442	99.2	882	23	AAS54848	Staphylococcus aur
3	1436	98.8	882	21	AAA92031	Staphylococcus aur
4	1415	97.4	864	23	AAS51602	Staphylococcus aur
5	1362	93.7	1893	18	AAV74466	Staphylococcus aur
6	1015	69.9	909	24	ABN90871	Staphylococcus epi
7	1015	69.9	3444	22	AAH54780	S. epidermidis gen
8	1015	69.9	4045	22	AAH54565	S. epidermidis gen
9	943	64.9	783	22	AAH52415	S. epidermidis ope
10	702.5	48.3	2944528	24	ABA03041	Listeria monocytog
11	695.5	47.9	5635	24	ABO71010	Listeria monocytog
12	690	47.5	413	23	AAS50245	Staphylococcus aur
13	686.5	47.2	495269	24	ABQ67195	Listeria innocua c
14	686.5	47.2	3011208	24	ABQ69245	Listeria innocua c
15	657	45.2	894	17	AAT40228	Mutant farnesylidip
16	656	45.1	894	17	AAT40229	Native farnesylidip
17	654	45.0	894	21	AAC63809	Bacillus stearothe
18	653	44.9	894	17	AAT40226	Mutant farnesylidip
19	650	44.7	894	17	AAT40227	Mutant farnesylidip
20	650	44.7	894	19	AAV38455	DNA encoding farne
21	650	44.7	894	19	AAV18516	Bacillus stearothe
22	650	44.7	894	20	AAK86781	Farnesyl diphospha
23	650	44.7	894	24	ABK96797	B. steatothermophi
24	647	44.5	894	14	AAQ39243	FFS DNA. Bacillus
25	645	44.4	894	17	AAT40225	Mutant farnesylidip
26	644	44.3	894	20	AAK86780	Geranyl diphosphat
27	642	44.2	894	24	ABK73724	B. steatothermophi
28	635	43.7	811	24	ABK73724	Bacillus lichenifo
29	596	41.0	882	23	AAS52896	Enterococcus faeca
30	587.5	40.4	7528	20	AAK12992	Enterococcus faeca
31	570	39.2	337	23	AAS50698	Staphylococcus aur
32	551.5	38.0	870	24	ABN6185	Streptococcus poly
33	551	37.9	876	21	AAA94607	Farnesyl diphospha
34	549	37.8	876	23	AAS55862	Streptococcus pneu
35	549	37.8	10011	19	AAV52209	Streptococcus pneu
36	547	37.6	876	23	AAS55538	Streptococcus pneu
37	540.5	37.2	2365589	24	ABA90521	Genomic sequence o
38	520.5	35.8	885	21	AAA13985	Taxus cuspidata ge
39	520.5	35.8	1179	21	AAA13994	Taxus GPP synthas
40	520.5	35.8	1179	21	AAA13995	Taxus GPP synthas
41	520.5	35.8	1179	21	AAA13996	Taxus GPP synthas
42	520.5	35.8	1179	21	AAA13997	Taxus GPP synthas
43	520.5	35.8	1179	21	AAA13998	Taxus GPP synthas
44	520.5	35.8	1179	21	AAA13999	Taxus GPP synthas
45	520.5	35.8	1179	21	AAA14000	Taxus GPP synthas

ALIGNMENTS

RESULT 1

AAS00821

ID AAS00821 standard; DNA; 861 BP.

XX AAS00821;

AC AAS00821;

DT 04-JUL-2001 (first entry)

XX S. aureus HGS072 encoding Farnesyl diphosphatesynthase, Ispa.

XX Farnesyl diphosphatesynthase; Ispa; immunogen; vaccine; antibody;

XX wound infection; cellulitis; burn infection; eyelid infection;

XX food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;

XX skin infection; scalded skin syndrome; toxic epidermal necrosis;

XX Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;

XX HGS072; ds.

XX Staphylococcus aureus.

OS

	361	GAACCTTATTTC	AAGTGATGATAGATTAACTGTAGTAAGCAATAAATAAAGTTCTACAACGG	420
Ddb				
OY	141	LeuSerIleAlaSerGlyHisValGlyMetValcylGlyGlnMetLeuAspMetGlnSer	160	
Ddb	421	CTGTCAATAGCAGTGGTCATGTTGGAATGGTCGGCGTCAAATGTTAGATATGC	480	
OY	161	GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla	180	
Ddb	481	GAAGGCCAACCAATTGATCTTGAACCTTTGGAANTGATACACAAAAACAAACAGGAGCA	540	
OY	181	LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValaspAspThrThrLys	200	
Ddb	541	CTATTAAACATTTCGGGTATGAGTGCAGCAGATATCCGTAATGTCGATGATGCAACTAAA	600	
OY	201	GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu	220	
Ddb	601	GAACATTTAGAAGTTATAGTTATCATCTTTAGGTATGATGTTTCAGATTAAAGATGATTA	660	
OY	221	LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlyLysSerAspLeuGluAsn	240	
Ddb	661	TTAGACTGCTATGGTGATGAAGCGCAAGTTTAGTAAAAAAGTGGCGACGATCTTGA	720	
OY	241	AsnLysSerThrTyrValSerLeuLeuGlyLysaspGlyAlaGluaspLysLeuThrTyr	260	
Ddb	721	AATAAAGTACATACCTGAGTTTATTAGGAAAAAGATGGCGAGAGATAAATTCAC	780	
OY	261	HisArgAspAlaAlaValaspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis	280	
Ddb	781	CATAGAGCGCACGTCGGTGAATGAATCAATTCGAATTCGACCAATTCATACAAACAC	840	
OY	281	LeuLeuGluIleValaspLeu	287	
Ddb	841	TTATTAGAATCGTTGATTTA	861	
	RESULT 3			
	AAA92031			
	ID AAA92031 standard; DNA; 882 BP.			
XX	AAA92031;			
XX	AC AC			
XX	DT DT			
XX	12-JAN-2001 (first entry)			
DE	Staphylococcus aureus ispA coding sequence.			
XX	IspA; bacterial disease; respiratory tract infection;			
KW	gastrointestinal infection; cardiac infection; Helicobacter pylori;			
KW	stomach cancer; stomach ulcer; gastritis; ds.			
XX	OS Staphylococcus aureus.			
XX	Key Location/Qualifiers			
FH	1..882			
FT	/tag= a			
FT	/product= "IspA"			
XX	US6107058-A.			
PN	22-AUG-2000.			
PD	26-MAR-1999; 99US-0276873.			
XX	26-MAR-1999; 99US-0276873.			
PF	(SMIK) SMITHKLINE BEECHAM CORP.			
PR	Gwynn M, Wilding EI;			
XX	WPI; 2000-578535/54.			
DR	P-PSDB; AAB23333.			
XX	Novel farnesyl diphosphate synthase polynucleotide from staphylococcus			
PT	aureus useful for diagnosis and treatment of bacterial infections and			
PT	as hybridization probe for isolating genomic clones -			
PT				

XX PS Claim 1: column 1-4; 15pp; English.
 CC The present sequence is the coding sequence for the Staphylococcus aureus
 CC IspA protein. This gene and the protein it encodes can be used in many
 CC research assays, as well as treatments for bacterial diseases such as
 CC infections of the respiratory tract (including otitis media, bacterial
 CC tracheitis, acute epiglottitis, thyroiditis, empyema and lung abscesses),
 CC cardiac infections such as infective endocarditis, gastrointestinal
 CC infections including secretory diarrhoea, splenic abscesses and
 CC retroperitoneal abscesses, CNS infections such as cerebral abscesses, eye
 CC infections (including blepharitis, conjunctivitis, keratitis,
 CC endophthalmitis, preseptal and orbital cellulitis and dacryocystitis),
 CC kidney and urinary tract infections such as epididymitis, intrarenal and
 CC perinephric abscesses and toxic shock syndrome, skin diseases (including
 CC impetigo, folliculitis, cutaneous abscesses, wound infection and
 CC bacterial myositis), bone and joint infections such as septic
 CC arthritis and osteomyelitis, septic thrombophlebitis, food poisoning and
 CC scalded skin syndrome. In addition, they can be used to treat
 CC diseases caused by Helicobacter pylori, including stomach cancer, stomach
 CC ulcers and gastritis.

SQ Sequence 882 BP; 328 A; 118 C; 186 G; 250 T; 0 other;

Alignment Scores:

Pred. No.: 1.23e-145 Length: 882
 Score: 1436.00 Matches: 284
 Percent Similarity: 99.30% Conservative: 1
 Best Local Similarity: 98.95% Mismatches: 2
 Query Match: 98.83% Indels: 0
 DB: 21 Gaps: 0

US-09-925-637-64 (1-287) x AAA92031 (1-882)

QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnGlnLeuSerValAla 20
 DB 1 ATGCGGAACTCCGATGAAATAAATAAGTAGAAGTCAATTAATGATTCGGTGGC 60
 QY 21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
 DB 61 ATAATAAATCACTAATGATCTAGTACAGTAGAAGAAAGCATGTTGTTATTAATGCT 120
 QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 DB 121 GGAGGTAAACCGCATCCGACGAGTCTGTATTACTCCTTACATTAATCAATACCGAG 180
 QY 61 TyrGluLeuGluMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
 DB 181 TATGAGTTAGGTGTGAGAGCGCAATGCTAGAAATGATTCATACATATTCACCTATT 240
 QY 81 HisAspLeuProAlaMetAspAsnAspTyrArgGlyLysLeuThrAsnHis 100
 DB 241 CATGATGACCTACCGCATGATGATATGATATCGAGGAAGAAATTAACAATCAT 300
 QY 101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
 DB 301 AAGATATATGCGTGGAGTGGATGATATAGCAGGTGATGCTTATTAACTAAAGCATTT 360
 QY 121 GluLeuIleSerSerAspArgLeuThrAspGluValLysLysValLeuLeuArg 140
 DB 361 GAACATTATTCAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSer 160
 DB 421 CTGTCAATGAAGTGGTTCATGTTGGAATGGTGGGGTCAAAATGTTAGATATGCAAGC 480
 QY 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
 DB 481 GAAGGCCAACCAATGATCTTGAACCTTTGGAATGATACACAAACAAACAGGAGCA 540
 QY 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspThrThrLys 200
 DB 541 TTATTNACNTTTCGGGTTATGAGTGCAGCATATCGCTAATGTCGATGATGCAACTAAA 600

QY 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetMetPheGlnIleLysAspLeu 220
 DB 601 GAACATTTAGAAAGTTACTTATCATTTAGGTATGATGTTCCAGATTAAGATGATTTA 660
 QY 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsn 240
 DB 661 TTAGACTGCTATGCTGATGAAGCAAAAGTTAGGTAAAAAGTGGCAGCGATCTTGAAT 720
 QY 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
 DB 721 AATAAAGTACGTACGTGAGTTTATTAGGAAAGATGGCCAGAAATAAATGACTTAT 780
 QY 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
 DB 781 CATAGAGCGCAGCAGTGGATGAACTAACCAATGATGAACAATTCATACAAACAC 840
 QY 281 LeuLeuGluIleValAspLeu 287
 DB 841 TTATTAGAAATCGTTGATTTA 861
 RESULT 4
 AAS51602
 ID AAS51602 standard; DNA; 864 BP.
 XX
 AC AAS51602;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus DNA for cellular proliferation protein #19.
 XX
 KW Antisense; ds; prokaryotic cellular proliferation gene;
 XX antibiotic; antibacterial; drug design.
 OS Staphylococcus aureus.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR P-PSDB; AAU33743.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Claim 27; Seq ID No 4184; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.

Sequence 864 BP; 323 A; 112 C; 182 G; 247 T; 0 other;

Pred. No.:	2,22e-143	Length:	864
Score:	1415.00	Matches:	280
Percent Similarity:	99.29%	Conservative:	0
Best Local Similarity:	99.29%	Mismatches:	2
Query Match:	97.38%	Indels:	0
DB:	23	Gaps:	0

QY	6	MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal	25
Db	1	ATGAATAAATAATAGATGAAGTCAATAAATGAATTATCGGTGCGATAAATAAATCAAGT	60
QY	26	MetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle	45
Db	61	ATGGATACTACGTAGAGAAGATGTTGTTATTCATTAATGCTGGAGGTAAACCGCATC	120
QY	46	ArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet	65
Db	121	CGACCCAGTCGTATTACTCACTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG	180
QY	66	LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro	85
Db	181	AAAGCGCAATTCGACTAGAAATGATTCATACATATTCACCTTATTCATGATGACCTACCA	240
QY	86	AlaMetAspAsnAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu	105
Db	241	CGCATGGATTAATGATTATTCGAGGAGGGAATTAACAATCATAAAGTATATGTGTAG	300
QY	106	TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer	125
Db	301	TGGACTGGCATATTAGCAGGTGATGCTTTATTAACATAAGCAATTTGAACCTATTTCAGT	360
QY	126	AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer	145
Db	361	GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAAAGGCTGTCAATAGCAAGT	420
Db	146	GlyHisValGlyMetValGlyGlyClnMetLeuAspMetGlnSerGluGlyGlnProIle	165
Db	421	GGTCATGTTGGTAATGGTCGGCGGTCAAAATGTTAGATGATGCAAGCGAAGGCCAACCAAT	480
QY	166	AspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAla	185
Db	481	GATCTTGAACTTGGAAATGATACACAAAAACAAACAGGAGCAGCTATTAAACATTGGC	540
QY	186	ValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSer	205
Db	541	GTTATGAGTCGACGAGATATCGCTAATGTCGATGATGCACTAAAGAACATTTAGAAGT	600
QY	206	TyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGly	225
Db	601	TATAGTTATCATTTAGGTATCATGTTTCAGATTAAAGATGATTTATTAGACTGCTATGGT	660
QY	226	AspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyr	245
Db	661	GATGAAGCGAAGTTAGTAAAAAGTGGCGAGCGATCTTGAAATAATAAAATACATAC	720
QY	246	ValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAla	265
Db	721	GTGAGTTATTAGGAAAAAGATGGCGCAAGATTAATTTGACTTATCATAGACGACGAGA	780

```
XX SQ Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;
Alignment Scores:
Pred. No.: 3,266-137 Length: 1893
Score: 1362.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.74% Indels: 0
DB: 18 Gaps: 0

US-09-925-637-64 (1-287) x AAV74466 (1-1893)
QY 19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeu 38
Db 1260 GTTGGCATAAATAACAGTATGATGATCAGTAGAAGAAAGTATGTTGATTTCATTA 1201
QY 39 AsnAlaGlyClyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsn 58
Db 1200 AATGCTGGAGGTAACCGCATCCGACCATGCTGTTATTACTCATTAGATTCACTAAAT 1141
QY 59 ThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSer 78
Db 1140 ACCGACTATGAGTATGATGATGAGAGCGCAATGCTAGAAATGATTCATCATATTC 1081
QY 79 LeuIleHisAspAspLeuProAlaMetAspAsnAspTyrArgArgGlyLysLeuThr 98
Db 1080 CTTATTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
QY 99 AsnHisLysValTyrGlyGluThrPheAlaIleLeuAlaGlyAspAlaLeuThrLys 118
Db 1020 AATCATAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 961
QY 119 AlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysValLeu 138
Db 960 GCATTGGAATTTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 901
QY 139 GlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuAspMet 158
Db 900 CAACGGCTGCAATAGCAAGTGTCTATGTTGAATGTCGCGGTCAAAATGTTAGATG 841
QY 159 GlnSerGluGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThr 178
Db 840 CAAGCGAAGGCAACCAATGATCTGAACTTTGGAATGATACACAAACAAACAA 781
QY 179 GlyAlaLeuLeuThrPheAlaValMetSerAlaIleAlaAspIleAlaAsnValAspThr 198
Db 780 GGAGCATTATTAACCTTTTCGGGTATGATGATGATGATGATGATGATGATGATGAT 721
QY 199 ThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAsp 218
Db 720 ACTAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAG 661
QY 219 AspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeu 238
Db 660 GATTATTATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
QY 239 GluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyValAlaGluAspLysLeu 258
Db 600 GAAATAATAAAGTACGTACGTGAGTTTATTAGGGAAGATGGCCGACAGATTAATG 541
QY 259 ThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThr 278
Db 540 ACTTATCATAGAGCAGCAGCAGTGGTGAATCAACCAAAATGATGACAAATTCATACA 481
QY 279 LysHisLeuLeuGluIleValAspLeu 287
Db 480 AAACACTTATTAGAAATCGTTGATTTA 454

RESULT 6
ABN90871
ID ABN90871 standard; DNA; 909 BP.
XX
```

```
AC ABN90871;
XX 24-JUL-2002 (first entry)
DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:334.
XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX Staphylococcus epidermidis.
XX OS Staphylococcus epidermidis.
XX PN US6380370-B1.
XX 30-APR-2002.
XX 13-AUG-1998; 98US-0134001.
XX 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA Doucette-Stamm LA, Bush D;
PI WPI; 2002-381255/41.
XX P-PSDB; ABP38326.
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX Disclosure; SEQ ID 334; 267pp; English.
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX SQ Sequence 909 BP; 361 A; 106 C; 165 G; 277 T; 0 other;

Alignment Scores:
Pred. No.: 3,61e-100 Length: 909
Score: 1015.00 Matches: 197
Percent Similarity: 82.93% Conservative: 41
Best Local Similarity: 68.64% Mismatches: 49
Query Match: 69.86% Indels: 0
DB: 24 Gaps: 0

US-09-925-637-64 (1-287) x ABN90871 (1-909)
QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
Db 28 ATGAGAAACTACAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 87
QY 21 IleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAla 40
Db 88 ATACAATCATCACCATTAAATACTTAATTTAGAAGAAGTATGAATATTCATTAAATGCT 147
QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 148 GGTGTTAAAGAAATCAGACCATCATATTATTATTAACTAAATGCTTAACTAAAGAT 207
QY 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
Db 208 TATCAACAGGACTAAATAGTCTTTAGCATTTGGAAATGATTCATATCTTCTTTAAT 267
QY 81 HisAspLeuProAlaMetAspAsnAspTyrArgArgGlyLysLeuThrAsnHis 100
```


QY	181	LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys	200
Db	1306	TTACTAAATTTTTCAGTTATGGCTGGGTAGACATTGCTCAAGTAGAACAAAATATTGCT	1365
QY	201	GUHHisLeuSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu	220
Db	1366	AGAANTTAGATGAATTTAGTCATCATTTTAGGAATGATGTTTCAAATTTAAAGATGATTTA	1425
QY	221	LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn	240
Db	1426	CTGGATGTGATGGTATGAATCAAACTGGCAGAAAAAGTAGCAGTATAGTAAAT	1485
QY	241	AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr	260
Db	1486	CATAAAAGTACTTATGTTCTTTACTTGGAAAAAGGAGCAGAGAAAAGTTAAACAAT	1545
QY	261	HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis	280
Db	1546	CATCAATATCTTGCTATGAACCTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA	1605
Db	281	LeuLeuGluIleValAspLeu	287
Db	1606	TTAAGTGATATGATGATTTA	1626
RESULT 8			
AAH54565/c			
ID	AAH54565	standard; DNA; 4045 BP.	
XX	AAH54565;		
AC	XX		
XX	XX		
DT	XX	03-SEP-2001 (first entry)	
XX	XX		
DE	XX	S. epidermidis genomic polynucleotide sequence SEQ ID NO:3929.	
XX	XX	Staphylococcus epidermidis SR1 strain; infection.; diagnosis;	
KW	XX	vaccination; endocarditis; ds.	
KX	XX		
OS	XX	Staphylococcus epidermidis.	
PN	XX	WO200134809-A2.	
PD	XX	17-MAY-2001.	
PF	XX	09-NOV-2000; 2000WO-US30782.	
PP	XX	09-NOV-1999; 99US-0164258.	
PR	XX	(GLAX) GLAXO GROUP LTD.	
XX	XX	Kimmerly WJ;	
XX	XX	WPI; 2001-316495/33.	
PT	XX	Nucleic acids encoding polypeptides from staphylococcus epidermidis,	
PT	XX	useful for vaccinating against infections, e.g. endocarditis -	
PS	XX	Claim 8; Page 1586-1588; 2188pp; English.	
PS	XX		
CC	XX	AAH5304 to AAH5370 represent nucleic acids (I) encoding polypeptides	
CC	XX	(II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.	
CC	XX	(I) and (II) can have antibacterial activity and therefore can be used	
CC	XX	in vaccination. The nucleic acids (I) may be used to produce the	
CC	XX	S. epidermidis polypeptides (II) via the production of vectors	
CC	XX	containing them which are used to produce hosts cells which express the	
CC	XX	polypeptides. The polypeptides (II) (and/or nucleic acids) may then be	
CC	XX	used to vaccinate subjects and to raise antibodies against the bacteria.	
CC	XX	The polypeptides may also be used to assay for other inhibitors of their	
CC	XX	activity and therefore identify compounds that may be used for the	
CC	XX	treatment of S. epidermidis infections, e.g. endocarditis. AAH53071 to	
CC	XX	AAH5090 represent specifically claimed S. epidermidis genomic DNA	
CC	XX	polynucleotide sequences from the present invention. AAH53091 to	
CC	XX	AAH5098 represent oligonucleotide sequences and primers which are used	
CC	XX	in the exemplification of the present invention.	

QY 281 LeuLeuGluLleValAspLeu 287
 DB 1473 TTAAGTGTATGTAGATTGA 1453

RESULT 9
 AAH52415
 ID AAH52415 standard; DNA; 783 BP.
 XX
 AC AAH52415;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:223.
 XX
 KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX
 OS vaccination; endocarditis; ds.
 XX
 OS Staphylococcus epidermidis.
 XX
 WO200134809-A2.
 17-MAY-2001.
 09-NOV-2000; 2000WO-US30782.
 09-NOV-1999; 99US-0164258.
 (GLAXO) GLAXO GROUP LTD.
 Kimmery WJ;
 WPI; 2001-316495/33.
 P-ESDB; AAG81565.

Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 useful for vaccinating against infections, e.g. endocarditis -
 Claim 8; Page 103; 218pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 (I) and (II) can have antibacterial activity and therefore can be used
 in vaccination. The nucleic acids (I) may be used to produce the
 S. epidermidis polypeptides (II) via the production of vectors
 containing them which are used to produce hosts cells which express the
 polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 used to vaccinate subjects and to raise antibodies against the bacteria.
 The polypeptides may also be used to assay for other inhibitors of their
 activity and therefore identify compounds that may be used for the
 treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 AAH55090 represent specifically claimed S. epidermidis genomic DNA
 polynucleotide sequences from the present invention. AAH55091 to
 AAH55098 represent oligonucleotide sequences and primers which are used
 in the exemplification of the present invention.
 N.B. The present invention specifically claims all the polynucleotide
 sequences given in the sequence listing of the present specification,
 however the sequence listing only goes up to SEQ ID NO:4454 so even
 though the sequences are given in the disclosure for SEQ ID NO:4472,
 no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 783 BP; 297 A; 93 C; 147 G; 246 T; 0 other;

Alignment Scores:
 Pred. No.: Length: 783
 Score: 943.00 Matches: 181
 Percent Similarity: 85.43% Conservativeness: 36
 Best Local Similarity: 71.26% Mismatches: 37
 Query Match: 64.90% Indels: 0
 DB: 22 Gaps: 0

US-09-925-637-64 (1-287) x AAH52415 (1-783)

QY 34 MetLeuTyrSerLeuAsnAlaGlyLysArgIleArgProValLeuLeuLeuLeuThr 53
 DB 1 ATGAATATTCATTAAATGCTGGTGTAAAGAAATCAGACCAGTCATCTATTATTATTAACA 60
 QY 54 LeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMet 73
 DB 61 CTAATAATGCTTACAAAGATTATCAACAGGACTTAATAGTGCCTTTAGCATTTGAAATG 120
 QY 74 IleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspPyrArg 93
 DB 121 ATTCATATCTTATCTTTAATTCATGATCATGATTTACCAGCAATGGATAATGACGATTACCGT 180
 QY 94 ArgGlyLysLeuThrAsnHisLysValTyrGlyGluTyrThrAlaIleLeuAlaGlyAsp 113
 DB 181 AGAGGAAATTAACAAATCAATAAGTTTATGGTGAATGGAAGCCATCTTTGCTGGTGTAT 240
 QY 114 AlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluVal 133
 DB 241 GCATTATTAACAAAGCTTTTGAATTAGTTTCTAATGATACCTACCTTTGAAGATAGTGTG 300
 QY 134 LysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGly 153
 DB 301 AAAGTAAGTATTATATAAAGAACTTTCAAAAGCAAGTGCACATTTGGGAATGGTGGTGGC 360
 QY 154 GlnMetLeuAspMetGlnSerGluGlnProIleAspLeuGluThrLeuGluMetIle 173
 DB 361 CAAGCGCTTGATATGGAAGTGAAGGGAAGTCAATTCGTTTGAAGAACTTTAGAAATCAATT 420
 QY 174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAla 193
 DB 421 CATGAACATAAGACAGCGCTTTTACTAAATTTTTCAGTATAGCTGCGGTAGACATTCCT 480
 QY 194 AsnValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMet 213
 DB 481 CAAGTAGAACAAATATTTGCTAAGAAATTTAGATGAATTTAGTCATCAATTTAGGAATGATG 540
 QY 214 PheGlnIleLysAspAspLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLys 233
 DB 541 TTTCAATTAAGATGATTTTACTGATGTGTGATGATGAATCAAACTTTGGCAAAAAA 600
 QY 234 ValGlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGly 253
 DB 601 GTAGCAGTGTATAGTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 660
 QY 254 AlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
 DB 661 GCAGAGAAAGAAAGTTAAACAATCATCAATATCTGCTATGCTGAAGCTGCTTAATCAATTTCT 720
 QY 274 GluGlnPheAsnThrLysHisLeuLeuGluLleValAspLeu 287
 DB 721 GATCAATATGATCTTCTGAATTAAGTATGATATTGTAGATTGA 762

RESULT 10
 ABA03041
 ID ABA03041 standard; DNA; 2944528 BP.
 XX
 AC ABA03041;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Listeria monocytogenes EGD-e genome sequence.
 XX
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 XX
 OS vitamin B12; bacterial infection; disease; ds.
 XX
 OS Listeria monocytogenes.
 XX
 PN WO200177335-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-FR01118.
 XX

```
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP ) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides
XX
XX Claim 1: SEQ ID No 1; 192pp; French.
XX
XX The present sequence is the genome sequence of Listeria monocytogenes
CC EGD-e. This sequence and fragments of this sequence are useful for
CC selecting probes and primers for detecting genes in L. monocytogenes and
CC related organisms, and to study genetic polymorphisms and other genomes.
CC Proteins (ABB47297-ABB50149) expressed from the present sequence are
CC useful for raising specific antibodies, identification of L.
CC Listeria monocytogenes and related organisms, and for biosynthesis and
CC biodegradation, especially biosynthesis of Vitamin B12. This sequence and
CC proteins encoded by it are also useful for selecting compounds that
CC regulate gene expression and cell replication and modulate L.
CC monocytogenes-related diseases. In addition, this sequence and
CC proteins encoded by it are useful in pharmaceutical and vaccine compositions for
CC the treatment or prevention of infections by L. monocytogenes and related
CC organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 2944528 BP; 91420 A; 563301 C; 555061 G; 911964 T; 0 other;

Alignment Scores:
Pred. No.: 7.57e-62 Length: 2944528
Score: 702.50 Matches: 148
Percent Similarity: 69.29% Conservative: 46
Best Local Similarity: 52.86% Mismatches: 81
Query Match: 48.35% Indels: 5
Gaps: 2

09-925-637-64 (1-287) x ABA03041 (1-2944528)
QY 8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
DB 1385970 AAAGTGATGATGAG-----TCGCTTTTAAAGAAATAACGCGGAAATATCGAA 1386020
QY 28 ThrGluLeuGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
DB 1386021 CCTAGACAAAGAGTCCATGTTATATCTATTCAAGCAGCGCGGAAAGGATTCGTCCA 1386080
QY 48 ValLeuLeuLeuLeuLeuLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
DB 1386081 ATGCTAGTTTTCCTACGCTTCAAGCTTAAAGTAAGTAATCCCGCTTTAGGTGTAAAAACC 1386140
QY 68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
DB 1386141 GCAACAGCGGTAGAAATGATTCACATACAGCTTATTCATGATGATCTACGAGCAATG 1386200
QY 88 AspAsnAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluThr 107
DB 1386201 GATAACGATGATTCGTGAGGCAAGTGGACTTAATCAAAAGTTTTTGGCGATGCAACT 1386260
QY 108 AlaIleAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAsp 127
DB 1386261 GCGATTTTGCAGGAGATGCTTTACTAACGCTCGCTTTTCTATTTTAGCTGAAGACGAT 1386320
QY 128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
DB 1386321 AATTATCTTTTGAGACACGCTTTGATTAACCAATAGTTTATAGTCGGGTGCA 1386380
QY 148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGlyGlyGlnProIleAspLeu 167
DB 1386381 GAAGGAATGTTGGTGGTCAACTTGCAGACTTGAAGCGGCAAAACAAGTACGCGTA 1386440
QY 168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
DB 1386441 GAAGAGTTATCATCCATTCATGCAGCAAAACGGGTGAATTTATTTATGCTGTAAAC 1386500
QY 188 SerAlaAlaAspIleAlaAsnValAspAspThrLysGluHisLeuGluSerTyrSer 207
DB 1386501 TCTGCAGCAAAATTCGGGAAGCTGATCCAGACAACAAACGCTTACCAATTTTTCGA 1386560
QY 208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
DB 1386561 GACAATATTCGGATTCGATTTCAAAATTTAGCGACGATATTTAGATGTAATTTGGTGATGAA 1386620
QY 228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
DB 1386621 ACGAAATGGGTAAAGACAGCGGCGCGCTTTCTGTAATAAAGTACCTATCCGGA 1386680
QY 248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
DB 1386681 TTACTACGCTTTCATGGGCAAAAGGCGCATTAATGAGCATGTTAGCATTTGCAAACTCA 1386740
QY 268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
DB 1386741 GCGCTTTCAGGCGCATGAT-----TTCGATGATGAAATTCCTTGAACACTTCTGATTTA 1386794

RESULT 11
ABQ71010
ID ABQ71010 standard; DNA; 5635 BP.
XX
XX AC ABQ71010;
XX
XX 29-AUG-2002 (first entry)
XX
XX Listeria monocytogenes 4b contig DNA sequence #952.
XX
XX Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
XX
XX Listeria monocytogenes 4b.
XX
XX WO200228891-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-FR03061.
XX
XX 04-OCT-2000; 2000FR-0012697.
XX
XX (INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Kunst F, Glaser P;
XX
XX WPI; 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
XX treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators
XX
XX Claim 14; SEQ ID 3823; 180pp; French.
XX
XX The present invention relates to nucleic acid sequences
XX (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
XX and primers for identification and/or detection of Listeria (e.g. as
```


Query Match: 47.49% Indels: 0
 DB: 23 Gaps: 0

US-09-925-637-64 (1-287) x AAS50245 (1-413)

QY 54 LeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaLeuGluMet 73
 DB 413 TTGATTTCACTAATACCGAGTAGTGGTTAGGTATGAAGACGCAATTGACACTAGAAATG 354

QY 74 IleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAspTyrArg 93
 DB 353 ATTCATACATATTCACCTATTCATGATGACCTACCGAGCGATGATGATGATATTCGA 294

QY 94 ArgGlyLysLeuThrAsnHisLysValTyrGlyLysThrAlaLeuAlaGlyAsp 113
 DB 293 CGAGGAAATTAACAATCAATAAGTATATGTTGAGTGAGTGGACTCGATATAGCAGGTGAT 234

QY 114 AlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluVal 133
 DB 233 GCTTTTATTAATAAGCATTTGAACCTATTTCAAGTGCATGATGATGATGATGATGATGATG 174

QY 134 LysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGly 153
 DB 173 AAAATAAAAGTTCTACACCGCTGCTCAATACGAAAGTGGTCTGTTGAAATGGTGGCGGT 114

QY 154 GlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
 DB 113 CAAATGTTAGATATGCAAGCGAAGGCAACCAATGATCTTCAAACTTTGGAAATGATA 54

QY 174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAla 190
 DB 53 CACAAAACAAAAACAGGAGCATTTATTAACCTTTTGGCGTTATGATGATGATGATGATGATG 3

RESULT 13
 ABQ67195/c
 ID ABQ67195 standard; DNA; 495269 BP.

XX ABQ67195;
 AC
 XX
 DT 29-AUG-2002 (first entry)
 XX
 DE Listeria innocua contig DNA sequence #8.
 XX
 XX Antibacterial; Listeria; food contamination; mutational analysis;
 KW infection; ds.
 XX
 OS Listeria innocua.
 WO200228891-A2.

PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-FR03061.
 XX
 PR 04-OCT-2000; 2000FR-0012697.
 XX
 PA (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Kunst F, Glaser P;
 PI
 XX WPI; 2002-332479/37.
 DR
 XX
 PT New genomic sequences from Listeria species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators
 XX
 PS Claim 5; SEQ ID 8; 180pp; French.
 XX

CC The present invention relates to nucleic acid sequences
 CC (ABQ67188-ABQ671212) from Listeria sp. The sequences are useful as probes
 CC and primers for identification and/or detection of Listeria (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of

CC gene expression. Proteins encoded by the nucleic acid sequences can be
 CC used to screen for compounds that modulate gene expression, replication
 CC and pathogenicity of Listeria (potential therapeutic agents), also for
 CC treating infections by Listeria, and are useful as immunogens in
 CC anti-Listeria vaccines.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 495269 BP; 159529 A; 88849 C; 93135 G; 153754 T; 2 other;

Alignment Scores:
 Pred. No.: 3.94e-61 Length: 495269
 Score: 686.50 Matches: 146
 Percent Similarity: 68.21% Conservative: 45
 Best Local Similarity: 52.14% Mismatches: 84
 Query Match: 47.25% Indels: 5
 DB: 24 Gaps: 2

US-09-925-637-64 (1-287) x ABQ67195 (1-495269)

QY 8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
 DB 366194 AAAGTACTTGATGAG-----TCGCTTTTAAAGAAATTAATATGCCCAATATCGAA 366144

QY 28 ThrGlnLeuGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
 DB 366143 CCTAAGTTGAAGAATCAATGTTATATTCAGTGCAGCTGCGGAAACCAATTCGTCCA 366084

QY 48 ValLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
 DB 366083 ATGCTTTGTTTTCGAACGCTTCAAGCGCTTAAATATTGAGCAATGCGCGGTTTAAACAA 366024

QY 68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
 DB 366023 GCTACGCGCACTGAAATGATTCATACGCTTAATTCACCATGATTATTCAGCAATG 365964

QY 88 AspAsnAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTyrThr 107
 DB 365963 GATGATGATGACTATCGACGCGGAAATGGACCAACCAATAAATATACGGTGATGCAACA 365904

QY 108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
 DB 365903 GCAATTTTAGCAGGAGATGCTTTGTAAACACTCGCTTTCTATTATTTAGCTGAAGATGA 365844

QY 128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
 DB 365843 AATCTATCTTTCGAAACGCGTATGCTTTAATTAATCAAAATAGTTATACCGTGGAGCA 365784

QY 148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGlyGlyGlnProIleAspLeu 167
 DB 365783 GAAGGCATGTTAGTGGCAACAACGCGATATGGAAGCTGAAATAAACAAGTCACATTA 365724

QY 168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
 DB 365723 GAAGAACTAGCATCAATCCACGCTCCCAAACTGGTGAACATATTAATTTTCGACGTAACC 365664

QY 188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
 DB 365663 TCAGCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAACGATTTACGAATTTTTCGA 365604

QY 208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuAspCysTyrGlyAspGlu 227
 DB 365603 GAAATATCGCATTTGTTTCAATATAGCAGCATATTTTACATGTAATTTGCGGATGA 365544

QY 228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
 DB 365543 ACAAAAATGGTAAAAAGACAGGGGTCGCGCTTTTCTGAATAAAGATACCTTCCCGGA 365484

QY 248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaValAsp 267
 DB 365483 TTACTCACACTCGAAGTCTTAACCGGCGCATTAATGACGATGTTTCAATTTGCAAGATCA 365424

Qy 268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
Db 365423 GCGCTTTTCAGGACATGAT-----TTCGACGATGAAATTCCTTAAACATTGCTGATTGA 365370

RESULT 14
ABO69245
ID ABO69245 standard; DNA; 3011208 BP.
XX
AC ABO69245;
XX
DT 29-AUG-2002 (first entry)
XX
DE Listeria innocua DNA sequence #684.
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
KW infection; ds.
XX
XX Listeria innocua.
OS
XX WO200228891-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-FR03061.
XX
XX 04-OCT-2000; 2000FR-0012697.
XX
XX (INSP) INST PASTEUR.
XX
XX (CNRS) CNRS CENT NAT RECH SCI.
XX
XX Kunst F, Glaser P;
XX
XX WPI; 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators -
XX
XX Claim 5; SEQ ID 2059; 180pp; French.
XX
XX The present invention relates to nucleic acid sequences
CC (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC and primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of
CC gene expression. Proteins encoded by the nucleic acid sequences can be
CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in
CC anti-Listeria vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;

Alignment Scores:
Pred. No.: 4,16e-60 Length: 3011208
Score: 686.50 Matches: 146
Percent Similarity: 68.21% Conservative: 45
Best Local Similarity: 52.14% Mismatches: 84
Query Match: 47.25% Indels: 5
DB: 24 Gaps: 2

US-09-925-637-64 (1-287) x ABO69245 (1-3011208)

Qy 8 LysLeuIleAspGluIleValAsnAsnGlnLeuSerValAlaIleAsnLysSerValMetAsp 27
Db 1392757 AAAGTACTGTATGAG-----TCGCTTTTAAAGAAATAATATCGCGCAATATCGAA 1392807

Qy 28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
Db 1392808 CCTAAGTTGAAGAATCAATGTTATATTCAGTCAGCTGCGGAAACGAATTCGTCCA 1392867

Qy 48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
Db 1392868 ATGCTTTTGGCAACCTTCAAGCTTAATATTGACCAATGCGGGTAAAAACA 1392927

Qy 68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
Db 1392928 GCTACGGCAGCTTGAAATGATTACATGACGTACAGCTTAATTCAGGATGATTACCAGCAATG 1392987

Qy 88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyLeuTyrPThr 107
Db 1392988 GATGATGATGACTATCGACGCGGGAATGGACGACCAATAAATATACGGTGATCAACA 1393047

Qy 108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
Db 1393048 GCAATTTTAGCAGGAGATGCTTTGTAACTCGCTTTTCTATTATTTAGCTGAAGATGAA 1393107

Qy 128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
Db 1393108 AATCTATCTTTCGAAACCGGTATTGCTTTAATTAATCAATAGTTATAGCAGTGGAGCA 1393167

Qy 148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGlyGlnProIleAspLeu 167
Db 1393168 GAAGCATGTTAGTGGGCAACAAGCCGATATGGAAGCTGAAATAAACAAGTCACATTA 1393227

Qy 168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
Db 1393228 GAAGAAGTATGATCATCAATCCAGCTCGCAAAACGTGTAACATTAATTTTCGACGTAACC 1393287

Qy 188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
Db 1393288 TCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAACGATTACGAATTTTTCGA 1393347

Qy 208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
Db 1393348 GAAATATCGCATTTGGTGTTCCTCAAAATAGCGACGATATTTTAGATGTAATTTGGCGATGAA 1393407

Qy 228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
Db 1393408 ACAAAAATGGGTAAAGACAGAGGGTTCGACGCTTTCTCTGAATAAAGTACCTATCCCGGA 1393467

Qy 248 LeuLeuGlyLysAspGlyAlaGluAspLysLysLeuThrTyrHisArgAspAlaAlaValAsp 267
Db 1393468 TTACTCACATCTCGAAGGTGCTTAACGGGCGCATTAATGAGCATGTTTCAATTTGCAAGTCA 1393527

Qy 268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
Db 1393528 GCGCTTTTCAGGACATGAT-----TTCGACGATGAAATTCCTTAAACTTGTCTGATTTA 1393581

RESULT 15
AAT40228
ID AAT40228 standard; DNA; 894 BP.
XX
AC AAT40228;
XX
DT 11-MAY-1997 (first entry)
XX
DE Mutant farnesylidiphosphate synthase (4).
XX
KW Farnesylidiphosphate; fpp; synthase; mutant; enzyme;
KW geranylgeranyl diphosphate; GGPP; ds.
XX
OS Bacillus stearothermophilus.
XX
XX EP733709-A2.
XX
XX 25-SEP-1996.
XX
XX 29-SEP-1995; 95EP-0115423.
XX
XX 14-FEB-1995; 95JP-0025253.
XX
XX (TOYT) TOYOTA JIDOSHA KK.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 20:32:16 ; Search time 59 Seconds
(without alignments)
1491.801 Million cell updates/sec

Title: US-09-925-637-64

Perfect score: 1453

Sequence: 1 MTNLPKMLIDEVNNELSA.....ELTQIDQFNKHLLEIVDL 287

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US0925637/runat_23052003_174818_23446/app_query.fasta_1.455
-DB=Issued Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0925637@cgn_1.1.40 @runat_23052003_174818_23446 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:**

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:**
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:**
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:**
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:**
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:**
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1436	98.8	882	3	US-09-276-873-1
2	1015	69.9	909	4	US-09-134-001C-334
3	657	45.2	894	1	US-08-534-910B-4
4	656	45.1	894	1	US-08-534-910B-5
5	654	45.0	893	1	US-08-333-321-1
6	653	44.9	894	1	US-08-534-910B-2
7	650	44.7	894	1	US-08-534-910B-3
8	650	44.7	894	3	US-08-886-466-1
9	650	44.7	894	4	US-09-475-304-1
10	650	44.7	894	4	US-09-101-126-4
11	650	44.7	894	4	US-09-367-528A-4
12	645	44.4	894	1	US-08-534-910B-1

13	644	44.3	894	4	US-09-367-528A-2	Sequence 2, Appli
14	551	37.9	876	3	US-09-275-742-1	Sequence 1, Appli
c 15	549	37.8	10011	4	US-08-961-527-76	Sequence 76, Appli
16	520.5	35.8	885	3	US-09-187-050-11	Sequence 11, Appli
17	520.5	35.8	1179	3	US-09-187-050-13	Sequence 13, Appli
18	520.5	35.8	1179	3	US-09-187-050-15	Sequence 15, Appli
19	520.5	35.8	1179	3	US-09-187-050-17	Sequence 17, Appli
20	520.5	35.8	1179	3	US-09-187-050-19	Sequence 19, Appli
21	520.5	35.8	1179	3	US-09-187-050-21	Sequence 21, Appli
22	520.5	35.8	1179	3	US-09-187-050-23	Sequence 23, Appli
23	520.5	35.8	1179	3	US-09-187-050-25	Sequence 25, Appli
24	520.5	35.8	1889	3	US-09-187-050-1	Sequence 1, Appli
25	465.5	32.0	1131	4	US-09-420-211-1	Sequence 1, Appli
26	438.5	30.2	486	3	US-09-217-609A-28	Sequence 28, Appli
27	438.5	30.2	486	3	US-08-873-235B-28	Sequence 28, Appli
28	376.5	25.9	909	1	US-07-783-705A-7	Sequence 7, Appli
29	376.5	25.9	6918	1	US-07-783-705A-13	Sequence 13, Appli
30	344	23.7	1284	4	US-08-858-207A-173	Sequence 173, App
31	340	23.4	1157	1	US-08-095-726-1	Sequence 1, Appli
32	340	23.4	1157	1	US-08-095-726-3	Sequence 3, Appli
33	340	23.4	1157	1	US-08-096-043-1	Sequence 1, Appli
34	340	23.4	1157	1	US-08-096-043-3	Sequence 3, Appli
35	340	23.4	1157	1	US-08-093-577-1	Sequence 1, Appli
36	340	23.4	1157	1	US-08-093-577-3	Sequence 3, Appli
37	340	23.4	1157	1	US-08-096-623A-1	Sequence 1, Appli
38	340	23.4	1157	1	US-08-096-623A-3	Sequence 3, Appli
39	332	22.8	993	1	US-08-705-377-2	Sequence 2, Appli
40	332	22.8	993	2	US-09-052-962-2	Sequence 2, Appli
41	332	22.8	993	2	US-09-053-068-2	Sequence 2, Appli
42	330	22.7	990	1	US-08-410-167A-1	Sequence 1, Appli
43	330	22.7	993	1	US-08-705-377-1	Sequence 1, Appli
44	330	22.7	993	2	US-09-052-962-1	Sequence 1, Appli
45	330	22.7	993	2	US-09-053-068-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-276-873-1
; Sequence 1, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwin Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: ISPA
; FILE REFERENCE: GM10208
; CURRENT APPLICATION NUMBER: US/09/276,873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-276-873-1

Alignment Scores:
Pred. No.: 4.34e-168 Length: 882
Score: 1436.00 Matches: 284
Percent Similarity: 99.30% Conservatives: 1
Best Local Similarity: 98.95% Mismatches: 2
Query Match: 98.83% Indels: 0
DB: 3 Gaps: 0

US-09-925-637-64 (1-287) x US-09-276-873-1 (1-882)

QY	1	MetThrAsnLeuPromMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
DB	1	ATGACGAATCTACCGATGAATAAATAGATAGATGAGTCAATGAATATCGGTGGG 60
QY	21	IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
DB	61	ATAAATAAATAGTATGATGCTACTGCTAGAGAAGCAATGTTGTATTATTAAATGCT 120

US-08-534-910B-4
; Sequence 4, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable
; OF SYNTHESIZING GERANYLGERANYLDIPHOSPHATE AND GENE CODING THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: IBM/Word Perfect 6.1 Windows
; CURRENT APPLICATION DATA:
; FILING DATE: 28-SEPT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US/08/534,910B
; FILING DATE: 14-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Toffenetti, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 77670/398
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)429-1776
; TELEFAX: (202)429-0796
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
; 08-534-910B-4
Alignment Scores:
Pred. No.: 6.79e-72 Length: 894
Score: 657.00 Matches: 133
Percent Similarity: 68.10% Conservative: 57
Best Local Similarity: 47.67% Mismatches: 83
Query Match: 45.22% Indels: 6
DB: 1 Gaps: 2
US-09-925-637-64 (1-287) x US-08-534-910B-4 (1-894)
QY 1 MetThrAsnLeuProMetAsnLysLeuLeuAspGluValAsnAsnGluLeuSerValAla 20
Db 1 ATCGCGCAGCTTCAGTTGAACAGTCTCAACGAGCAAAACAGCGGTGGAACAGCG 60
QY 21 IleAsnLysSerValMet-----AspThrGlnLeuGluGluSerMetLeuTyr 36
Db 61 CTCCTCCGTTATATAGCGCTTAGAGCGCGCGGAGCTGAAGAGGCGGTGCGCGT 120
QY 37 SerLeuAsnAlaGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSer 56
Db 121 TCATTGAGCGCGCGGCAACCAATCCGTCCTGCTGCTTCTGTCCACCGTTCGGCG 180

QY 57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
Db 181 CTCGGCAAGACCGCGCGCGGATTCGCGCTGCGCGATGAATGAATCATCATACG 240
QY 77 TyrSerLeuIleHisAspAspLeuProAlaMetAsnAspAspTyrArgArgGlyLys 96
Db 241 CACTCTTTGATCCATGATGATTCGCCGAGCATGACACACGATGATTTCGCGCGCGG 300
QY 97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeu 116
Db 301 CCGACGAAACCAAAAGTTCGCGGAGCGGATGCGCATCTTCGCGGCGGAGGTTCTG 360
QY 117 ThrLysAlaPheGluLeuIleSer-----SerAspArgLeuThrAspGluValLys 134
Db 361 ACGTACGCGTTTCAATTGATCACGAAATGACGATGAGCGCATCCCTCTCCGTCGG 420
QY 135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGln 154
Db 421 CTTGCGCTCATGCAACGCGTGGGAAAGCGCGGTCCGGAAGGATGGTCCGCGGT 480
QY 155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174
Db 481 GCAGCGCATATGGAAGGAGAGGGAACGCTGACGCTTCGGAGCTCGAATACATCAT 540
QY 175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaIleAspIleAlaAsn 194
Db 541 CGGCATAAAACCGGGAATGCTGCAATACAGCGTGCACGCGCGCGCTTGATCGCGG 600
QY 195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214
Db 601 GCTGATCCCGGCAACGCGGAGCTTGAGCAATTCGCGCCCATCTAGGCTTCGCTTT 660
QY 215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysVal 234
Db 661 CAAATTCGCGATGATATCTCGATATTGAAGGGGACAGAAATAATCGCAACCGGTC 720
QY 235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
Db 721 GGCAGCGCAACCAAGCAACAAAGCAGCATATCAGCGTTCGCTGCTGCGTTCGCGG 780
QY 255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
Db 781 AAGAAAGTTGACGTTCCATATCGAGCGCGCGGCGCATTTACGGAACGCGCGAC 837
RESULT 4
US-08-534-910B-5
; Sequence 5, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyoza
; APPLICANT: KOYAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable
; OF SYNTHESIZING GERANYLGERANYLDIPHOSPHATE AND GENE CODING THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: IBM/Word Perfect 6.1 Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/534,910B

```

QY 195 ValAspAspThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 234
Db   |||  |||:::  :::::  ||| ||| |||:::  |||
601 GCTGATGCCCGCAACGCGGAGCTTCAGCAATTCGCGCCCATCTAGCGCTTGCCTTT 660
QY 215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
661 CAAATTCCGGATGATATTCTCGATATTGAAGGGGCGACAGAAAATTCGCAAGCCGGTC 720
QY 235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
721 GCGAGGACCAAGCAACAACAACGACGATGATCCAGCGTGTCTGCTTGCCTTGCAGCGGCG 780
QY 255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
781 AAGGAAAAGTTGGCTTCATATCAGGCGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 5
US-08-333-321-1
; Sequence 1, Application US/0833321
; Patent No. 5786192
; GENERAL INFORMATION:
; APPLICANT: Obata, Shusei
; TITLE OF INVENTION: FARNESYL PYROPHOSPHATE SYNTHETASE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward W. Greason, Esq.
; STREET: 1 Broadway
; CITY: New York
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: No. 5786192epad, Windows 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,321
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/953,424
; FILING DATE: 29-SEP-1992
; APPLICATION NUMBER: JP 3-253788
; FILING DATE: 01-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Greason, Edward W.
; REGISTRATION NUMBER: 18,918
; REFERENCE/DOCKET NUMBER: 077670/00310
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 893 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
US-08-333-321-1

Alignment Scores:
Pred. No.: 1.59e-71 Length: 893
Score: 654.00 Matches: 133
Percent Similarity: 67.38% Conservative: 55
Best Local Similarity: 47.67% Mismatches: 85
Query Match: 45.01% Indels: 6
DB: 1 Gaps: 2

US-09-925-637-64 (1-287) x US-08-333-321-1 (1-893)

```


QY		135	IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln	154
Db		421	CTTCGGCTCATCGAACGGCTGGCAAGACGGCCGGTCCGAAGGGATGTGC CGCGGT CAG	480
QY		155	MetLeuAspMetGlnSerGluglyGlnProIleAspLeucluthrLeuGluMetIleHis	174
Db		481	GCAGCCGATATGGAAAGAGAGGGGAAAACGCTCAGCCTTCGGAGCTCGAATACATT CAT	540
QY	.	175	LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn	194
Db		541	CGGCATAAAACCGGAAAAATGCTCAATACAGGTGCACGCCGCCCTTGATCGCGCG	600
QY		195	ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe	214
Db		601	GCTGATGCCCGGCAACCGCGGAGCTTGAGCAATTCGCCCCCATCTAGCGCTTGCCCTT	660
QY		215	GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal	234
Db		661	CAAAATTCGGATGATATTCGATATTGAAGGGGCGAGAAGAAAAATCGCAAGCGCGTC	720
		235	GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla	254
Db		721	GGCAGCGACCAACAACAACAGCACGATCCAGCGTTGCTGCTGCTGCCGCGCG	780
QY		255	GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp	273
Db		781	AAGAAAAAGTTGGGCTTCCATATCGAGCGGCGCAGCGCATTTACGAAACGCCGAC	837

RESULT 7

US-08-534-910B-3
Sequence 3, Application US/08534910B
Patent No. 5766911
GENERAL INFORMATION:
APPLICANT: KOIKE, Ayumi
APPLICANT: OBATA, Shusei
APPLICANT: NISHINO, Tokuzo
APPLICANT: OHNUMA, Shinichi
APPLICANT: NAKAZAWA, Takeshi
APPLICANT: OGURA, Kyoza
APPLICANT: Koyama, Tanetoshi
TITLE OF INVENTION: Mutated Farnesylphosphate Synthase Capable
OF SYNTHESIZING GERANYLGERANYLDIPHOSPHATE AND GENE CODING THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, N.W., Suite. 600
CITY: Washington
STATE: DC
COUNTRY: U.S.
ZIP: 20036-5405
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.25" Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: IBM/Word Perfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/534,910B
FILING DATE: 28-SEPT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-25253
FILING DATE: 14-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 77670/398
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)429-1776
TELEFAX: (202)429-0796
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 base pairs
TYPE: nucleic acid

RESULT 8

US-08-886-466-1
; Sequence 1, Application US/08886466C
; Patent No. 6040165
; GENERAL INFORMATION:
; APPLICANT: Narita, Keishi
; APPLICANT: Ishida, Chika
; APPLICANT: Takeuchi, Yoshie
; APPLICANT: Ohto, Chikara
; APPLICANT: Ohnuma, Shinichi
; APPLICANT: Nishino, Tokuzo
; TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
; FILE REFERENCE: 77670/494
; CURRENT APPLICATION NUMBER: US/08/886,466C
; CURRENT FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: JP 8-191635
; EARLIER FILING DATE: 1996-07-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
LENGTH: 894
TYPE: DNA
ORGANISM: Bacillus stearothermophilus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(891)
US-08-886-466-1

Alignment Scores:
Pred. No.: 4,97e-71 Length: 894
Score: 650.00 Matches: 134
Percent Similarity: 67.78% Conservative: 49
Best Local Similarity: 49.63% Mismatches: 85
Query Match: 44.74% Indels: 2
DB: 3 Gaps: 1

US-09-925-637-64 (1-287) x US-08-886-466-1 (1-894)

QY 6 MetAsnLysLeuLeuValAsnGluLeuSerValAlaLeuSerVal 25
Db 28 CTCACGAGCAACAGCGCGTGGAAACAGCGCTCTCCGTTATATAGAGCGTTAGAA 87
QY 26 MetAspThrGlnLeuGluSerMetLeuThrLeuAsnAlaGlyLysArgile 45
Db 88 GGGCGCGGCAAGCTGAAAGCGATGCGGTACTCATTTGGAGCGCGGCAACGAATC 147
QY 46 ArgProValLeuLeuLeuThrLeuAspSerLeuThrGluLeuGlyMet 65
Db 148 CGTCGCTGCTCTCTGTCCACCGTTCGGCGCTCGCAACAGCCCGGCGTGGATTG 207
QY 66 LysSerAlaLeuAlaLeuGluMetIleHisThrYrSerLeuLeuHisAspLeuPro 85
Db 208 CCGCTGCGCTGCGCGATGAAATGATCCATAGTACTCTTTCATCATGATTTGCGG 267
QY 86 AlaMetAspAsnAspArgArgGlyLysLeuThrAsnHisLysValThrGlyGlu 105
Db 268 AGCATGGACAGATGATTTGCGCGCGCAAGCGAGCCATATAAGTGTTCGGCGAG 327
QY 106 TrpThrAlaLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuLeuSer 124
Db 328 GCGATGGCCACTTTCGGCGGGGACCGGTGTGTGACGTACGCTTCAATGATACCGAA 387
QY 125 ---SerAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIle 143
Db 388 ATCGACGATGAGCGCATCCCTCTTCGCTCGCGCTTCGGCTCATCGAAGCGGTGGCAA 447
QY 144 AlaSerGlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSerGlyGln 163
Db 448 GCGCGCGTCCGGAAGGATGTCGCGGTTCAGGCGAGCGGATATGGAAGGAGAGGAA 507
QY 164 ProIleAspLeuGluThrMetIleHisLysThrLysThrGlyAlaLeuLeuThr 183
Db 508 ACGGTGACGCTTTCGGAGCTCAATACATTCATCGCATATAAACCGGGAATGCTGCA 567

RESULT 9

US-09-475-304-1
; Sequence 1, Application US/09475304
; Patent No. 6225096
; GENERAL INFORMATION:
; APPLICANT: Narita, Keishi
; APPLICANT: Ishida, Chika
; APPLICANT: Takeuchi, Yoshie
; APPLICANT: Ohto, Chikara
; APPLICANT: Ohnuma, Shinichi
; APPLICANT: Nishino, Tokuzo
; TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
; FILE REFERENCE: 77670/494
; CURRENT APPLICATION NUMBER: US/09/475,304
; CURRENT FILING DATE: 1999-12-30
; EARLIER APPLICATION NUMBER: JP 8-191635
; EARLIER FILING DATE: 1996-07-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
LENGTH: 894
TYPE: DNA
ORGANISM: Bacillus stearothermophilus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(891)
US-09-475-304-1

Alignment Scores:
Pred. No.: 4,97e-71 Length: 894
Score: 650.00 Matches: 134
Percent Similarity: 67.78% Conservative: 49
Best Local Similarity: 49.63% Mismatches: 85
Query Match: 44.74% Indels: 2
DB: 4 Gaps: 1

US-09-925-637-64 (1-287) x US-09-475-304-1 (1-894)

QY 6 MetAsnLysLeuLeuValAsnGluLeuSerValAlaLeuSerVal 25
Db 28 CTCACGAGCAACAGCGCGTGGAAACAGCGCTCTCCGTTATATAGAGCGTTAGAA 87
QY 26 MetAspThrGlnLeuGluSerMetLeuThrLeuAsnAlaGlyLysArgile 45
Db 88 GGGCGCGGCAACAGCGCGTGGAAAGCGATGCGGTACTCATTTGGAGCGCGGCAACGAATC 147
QY 46 ArgProValLeuLeuLeuThrLeuAspSerLeuThrGluLeuGlyMet 65
Db 148 CGTCGCTGCTCTCTGTCCACCGTTCGGCGCTCGCAACAGCCCGGCGTGGATTG 207
QY 66 LysSerAlaLeuAlaLeuGluMetIleHisThrYrSerLeuLeuHisAspLeuPro 85

Db 208 CCCGTCGCCCTCGCGGATTTGAATGATCCATACGTACTCTTTGATCCATGATGATTTCCGG 267
Qy 86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
Db 268 AGCATGACAAACGATTTGCGCGCGGCAAGCGGACGAACCAATGAAGTGTTCGCGCGAG 327
Qy 106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer--- 124
Db 328 GCGATGCCCATCTTGGCGGGGAGGTTGTGTACGTACGCGTTCAATTCATCATCCGCAA 387
Qy 125 ---SerAspAspArgLeuThrAspGluValLysLysLysLysLysLysLysLysLysLys 143
Db 388 ATCGACGATGAGCGCATCCCTCTTCGTCGCGCTTCGCTCAFCGAAACGCGTGGCGAAA 447
Qy 144 AlaSerGlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSerGlyGln 163
Db 448 GCGCGCGTCCGGAAGGATGGTTCGCGGTTCAGCGACCGCATATGGAAGGAGAGGGGAAA 507
Qy 164 ProfileAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThr 183
Db 508 ACGTGTAGCGTTTCGGAGCTCGAATACATTCATCGGCATAAACCGGGAATGCTGCAA 567
Qy 184 PheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeu 203
Db 568 TACAGCGTGACGCGCGCGCTTGATTCGCGGCGCTGATCCCGGCAACCGCGGAGCTT 627
Qy 204 GluSerTyrSerTyrHisLeuGlyMetPheGlnIleLysAspAspLeuLeuAspCys 223
Db 628 GACGAATTCGCGCGCCATCTAGCGCTTTCGCTTCAATTCGCGATATTCGATATT 687
Qy 224 TyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsnAsnLysSer 243
Db 688 GAAGGGCAGCAAGAAAATTCGCAAGCGGTTCGCGACGACCAACCAACAAAGCG 747
Qy 244 ThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAsp 263
Db 748 ACGTATCCAGGTGCTGCTGCTGCGCGCGGCAAGGAAAAGTGGCGTTCATATCGAG 807

RESULT 10

US-09-101-126-4

; Sequence 4, Application US/09101126

; Patent No. 6316216

; GENERAL INFORMATION:

APPLICANT: OHTO, CHIKARA

APPLICANT: NAKANE, HIROYUKI

APPLICANT: NISHINO, TOKUZO

APPLICANT: OHNUMA, SHINICHI

APPLICANT: HIROOKA, KAZUTAKE

; TITLE OF INVENTION: MUTATED PRENYL DIPHOSPHATE SYNTHASES

; FILE REFERENCE: 77670/566

; CURRENT APPLICATION NUMBER: US/09/101,126

; CURRENT FILING DATE: 1999-04-27

; EARLIER APPLICATION NUMBER: PCT/JP97/03921

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: JP 8-307506

; EARLIER FILING DATE: 1996-11-05

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn ver. 2.0

; SEQ ID NO 4

; LENGTH: 894

; TYPE: DNA

; ORGANISM: Bacillus stearothermophilus

; FEATURE:

; OTHER INFORMATION: 256-276 is an Asp-rich coding domain

US-09-101-126-4

Alignment Scores:

Pred. No.: 4,97e-71

Score: 650.00

Length: 894

Matches: 134

Percent Similarity: 67.78% Conservative: 49
Best Local Similarity: 49.63% Mismatches: 85
Query Match: 44.74% Indels: 2
DB: 4 Gaps: 1

US-09-925-637-64 (1-287) x US-09-101-126-4 (1-894)

Qy 6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
Db 28 CTCAACGACAAACAAACAGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTGAA 87
Qy 26 MetAspThrGlnLeuGluGluSerMetLeuThrSerLeuAsnAlaGlyGlyLysArgIle 45
Db 88 GGGCCGCGCAAGTGAAGGCGATGGCTACTCATTTGGAGCGCGCGCAACGAAATC 147
Qy 46 ArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 65
Db 148 CGTCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207
Qy 66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
Db 208 CCCGTCGCCCTCGCGGATTTGAATGATCCATACGTACTCTTTGATCCATGATTTCCG 267
Qy 86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
Db 268 AGCATGACAAACGATTTGCGCGCGGCAAGCGGACGAACCAATGAAGTGTTCGCGCGAG 327
Qy 106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer--- 124
Db 328 GCGATGCCCATCTTGGCGGGGAGGTTGTGTACGTACTCTTTCAATTCATCATCCGAA 387
Qy 125 ---SerAspAspArgLeuThrAspGluValLysLysLysLysLysLysLysLysLys 143
Db 388 ATCGACGATGAGCGCATCCCTCTTCGTCGCGCTTCGCTCAFCGAAACGCGTGGCGAAA 447
Qy 144 AlaSerGlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSerGlyGln 163
Db 448 GCGCGCGTCCGGAAGGATGGTTCGCGGTTCAGCGACCGCATATGGAAGGAGAGGGGAAA 507
Qy 164 ProfileAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThr 183
Db 508 ACGTGTAGCGTTTCGGAGCTCGAATACATTCATCGGCATAAACCGGGAATGCTGCAA 567
Qy 184 PheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeu 203
Db 568 TACAGCGTGACGCGCGCGCTTGATTCGCGGCGCTGATCCCGGCAACCGCGGAGCTT 627
Qy 204 GluSerTyrSerTyrHisLeuGlyMetPheGlnIleLysAspAspLeuLeuAspCys 223
Db 628 GACGAATTCGCGCGCCATCTAGCGCTTTCGCTTCAATTCGCGATATTCGATATT 687
Qy 224 TyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsnAsnLysSer 243
Db 688 GAAGGGCAGCAAGAAAATTCGCAAGCGGTTCGCGACGACCAACCAACAAAGCG 747
Qy 244 ThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAsp 263
Db 748 ACGTATCCAGGTGCTGCTGCTGCTGCGCGCGGCAAGGAAAAGTGGCGTTCATATCGAG 807
Qy 264 AlaAlaValAspLeuLeuThrGlnIleAsp 273
Db 808 GCGGCGCAGCGCCATTTACGGAACGCGCGAC 837

RESULT 11

US-09-367-528A-4

; Sequence 4, Application US/09367528A

; Patent No. 6395525

; GENERAL INFORMATION:

APPLICANT: TOYOTA JIDOSHA KABUSHIKI KAISHA

; TITLE OF INVENTION: Geranyl Diphosphate Synthetase Gene

; FILE REFERENCE: PH-586

; CURRENT APPLICATION NUMBER: US/09/367,528A

; CURRENT FILING DATE: 1999-08-16

QY	1	MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla	20
		:::: ::::	
DB	1	ATGGCGCAGCTTTCAGTTGAACAGTTCTCAACGAGCAAAACAGCGGTGGAAACACGG	60
QY	21	IleAsnLysSerValMet-----AspThrGlnLeuGluGlnSerMetLeuTyr	36


```

; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: Ispa
; FILE REFERENCE: GMI0205
; CURRENT APPLICATION NUMBER: US/09/275,742
; CURRENT FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-375-742-1

```

Alignment Scores:		
Pred. No.:	8, 076-59	876
Score:	531.00	122
Percent Similarity:	64.8%	Conservative: 48
Best Local Similarity:	46.5%	Mismatches: 88
Query Match:	37.9%	Indels: 4
DB:	3	Gaps: 3

-09-925-637-64 (1-287) x US-09-275-742-1 (1-876)

QY	28	ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro	47
Db	73	TCTAGTTTACGGGAGCTGTGTTCTATTCTATTCTGCTGGCAAGCGTATTGCGCCT	132
QY	48	ValLeuLeuLeuThrLeuAspSerLeuAsn---ThrGluTyrGluLeuGlyMetLys	66
Db	133	TTTCTCTGTTTAAAGATTCTGGAAGCCTTCAGGTTTACCATAAACCTGCTCAGCGGCAG	192
QY	67	SerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspSerLeuProAla	86
Db	193	GTAGCTACTGCTTGGAGATGATTCATACAGGAGCTTGATTACAGATGACCTTCTCTGCT	252
QY	87	MetAspAsnAspTyrArgGlyLysLeuThrAsnHisLysValTyrGlyGluTyr	106
Db	253	ATGGATGACGATTATCGAAGGCGGTTAACCAATCACAAGAAATTCGTGGAAGCT	312
QY	107	ThrAlaIleLeuAlaGlyAspAlaLeuThrLysAlaPheGluLeuIleSerSerAsp	126
Db	313	ATGGCAATTGTCCTGGAGATGCCTTATTCCTTAGACCATGCTTGATTGGCGCAGCA	372
QY	127	AspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGly	146
Db	373	GAT---TTGCCAAGTCAGATCAGGCGGGCTTGATTGCCCAACTTATCCCTTGCTTCAGGT	429
QY	147	HisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp	166
QY	430	AGTCGGGTATGGTGGCAGGCGCAGTTTGGATATGAGGGCGGAACACCAAGCACTTGCTCT	489
QY	167	LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal	186
Db	490	CTGGAAGAACTTCAGACTATTTCATGCCCAATAAGACTGGGAAGTTACTAGCTATCCCTTC	549
QY	187	MetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyr	206
Db	550	CAAGCGGACGTATTATAGCTGAATTGTCACCTGAAATGCAGGTGAAGCTGAAAACTGTG	609
QY	207	SerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAsp	226
Db	610	GCTGAATTGATTGGACTTCTCTTTCAAGTCAGAGATGATCTACTGGATGTGACGGCTAGT	669
QY	227	GluAlaLysLeuGlyLysValGlySerAspLeuGluAsnAsnLysSerThrTyrVal	246
Db	670	TTTGGAGAAATCGGCACACACCTCAAAAGAGCTCGCAGGCGAAGAAATCAACCTTATCCT	729
QY	247	SerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaVal	266
Db	730	GCTTGTGGGCTTGGAGAGTCCATTGCGCTTTTGTAAACCAACCCTGGATCAAGCTAAT	789
QY	267	AspGluLeuThrGlnIleAspGluGln-----PheAsnThrLysHisLeuLeuGluIle	284

```

Db      790 GATAAATTAGACAGAAATGCCCCAGCAGCTTCCTTTGAAACAGAAATCGATTGCTAAAGTCTA 849
Qy      285 Valasp 286
      |||::
Db      850 GTAGAA 855

RESULT 15
US-08-961-527-76/c
: Sequence 76, Application US/08961527
: Patent No. 6420135
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 391
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,527
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 76:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1001 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-961-527-76

```

Alignment Scores:		
Pred. No.:	5,516-57	10011
Score:	549.00	Matches: 121
Percent Similarity:	65.27%	Conservative: 50
Best Local Similarity:	46.18%	Mismatches: 87
Query Match:	37.78%	Indels: 4
DB:	4	Gaps: 3

US-09-925-637-64 (1-287) x US-08-961-527-76 (1-10011)

[illegible]

Qy 48 ValLeuLeuLeuThrIfeuAspserLeuAse---ThrGlutyrGluLeuGlyMetLys 66
 ||||| ||||| ||||| ||||| ||||| ||||| : : :
Db 7746 TTTCTCTGTAGAACTTCTGAAGCCTTGCAAGTTACCATCAAACCTGCTCACGCAG 7687

Qy	67	SerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAla	86
Db	7686	GTAGCTACTGCCTTAGAGATGATTCATACAGGGAGCTTGATTCACGATGACCTTCCTGCT	7637

Qy	87	MetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp	106
		
		
		
		
Db	7626	ATGGATGATGACGATTATCGAAGAGGCGCGTTAACCAATCACACGAATAATCGGTGAAGCT	7567
		
		
		
		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 20:55:11 ; Search time 187 Seconds
(without alignments)
2069.416 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPKMLIDEVNNELSV.....ELTQIDKQFNKHLLEIVDL 287

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

earched: 845702 seqs, 674182571 residues

Total number of hits satisfying chosen parameters: 1691404

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame-p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09925637/runat_23052003_174819_23502/app_query.fasta_1.455
-DB=Published Applications_NA -QWTF=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09925637 -CGN_1_1_57 -runat_23052003_174819_23502
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFLOCK=100
-LANGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	1453	100.0	861	9	US-10-084-205-63
2	1453	100.0	861	10	US-09-925-637-63
3	1453	100.0	1893	7	US-08-781-986A-155
4	1442	99.2	882	10	US-09-815-242-8485
					Sequence 63, Appl
					Sequence 63, Appl
					Sequence 155, App
					Sequence 8485, Ap

5	1415	97.4	864	10	US-09-815-242-4184
6	635	47.5	413	10	US-09-815-242-2822
7	635	43.7	811	10	US-09-974-300-1015
8	596	41.0	882	10	US-09-815-242-6533
9	587.5	40.4	7528	10	US-09-070-977A-55
10	570	39.2	337	10	US-09-815-242-3275
11	549	37.8	876	10	US-09-815-242-9499
12	547	37.6	876	10	US-09-815-242-9175
13	516.5	35.5	891	9	US-09-941-947A-19
14	516.5	35.5	891	10	US-09-934-903-13
15	516.5	35.5	891	10	US-09-934-868-71
16	504.5	34.7	1470	9	US-10-108-915-25
17	497.5	34.2	888	10	US-09-815-242-7142
18	494.5	34.0	1441	9	US-10-108-915-21
19	487.5	33.6	1268	9	US-10-108-915-17
20	485.5	33.4	900	10	US-09-815-242-5972
21	474.5	32.7	888	10	US-09-815-242-7873
22	473	32.6	900	10	US-09-815-242-9986
23	465.5	32.0	1131	10	US-09-934-778-1
24	380.5	26.2	1062	9	US-10-108-915-15
25	376.5	25.9	912	9	US-09-941-947A-25
26	375	25.8	640681	10	US-09-790-988-1
27	373.5	25.7	912	10	US-09-815-242-7285
28	372.5	25.6	912	10	US-09-815-242-7450
29	356	24.5	973	9	US-10-108-915-13
30	352	24.2	536165	9	US-09-933-964-1
31	352	22.8	951	9	US-10-166-037-3
32	328	22.6	498	10	US-09-974-300-5453
33	328	22.6	973	9	US-10-108-915-11
34	321	22.1	498	10	US-09-974-300-5438
35	311.5	21.4	969	10	US-09-815-242-8505
36	311.5	21.4	1161	9	US-10-108-915-37
37	311	21.4	972	10	US-09-815-242-6224
38	310	21.3	972	10	US-09-815-242-9636
39	309.5	21.3	888	10	US-09-547-367-2
40	309.5	21.3	8625	9	US-09-920-923-1
41	309.5	21.3	11233	9	US-09-920-923-27
42	308.5	21.2	936	10	US-09-815-242-4260
43	297.5	20.5	1253	9	US-09-920-923-35
44	297	20.4	2000	10	US-09-887-576-115
45	294	20.2	990	10	US-09-815-242-7045

ALIGNMENTS

RESULT 1
US-10-084-205-63
; Sequence 43, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: No1, G11
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: P515P1
; CURRENT APPLICATION NUMBER: US/10/084, 205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOs: 74
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-084-205-63

Alignment Scores:
Pred. No.: 2, 1e-161
Score: 1453.00
Percent Similarity: 100.00
Best Local Similarity: 100.00
Query Match: 100.00
Length: 861
Matches: 287
Conservative: 0
Mismatches: 0
Indels: 0

```
DB: 9 Gaps: 0
US-09-925-637-64 (1-287) x US-10-084-205-63 (1-861)
QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
Db 1 ATGACGAATCTACCGATGAATTAATATAGATGAATCAATGAATATCGGTGGC 60
QY 21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
Db 61 ATAATAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 121 GGAGGTAACCGATCGACCGATCTGTTATTAATCAATCAATCAATCAATCAATCAAT 180
QY 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
Db 181 TATGAGTTAGGTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 81 HisAspAspLeuProAlaMetAspAspTyrArgArgGlyLysLeuLeuLeuLeuLeuLeu 100
Db 241 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuLeuLeuLeuLeu 120
Db 301 AAAGTATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 121 GluLeuLeuSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
Db 361 GAACCTTATTTCAAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlyGlyGlyGlyGlyGlyGly 160
Db 421 CTGTCAATAGCAAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisThrLysThrGlyAla 180
Db 481 GAAGGCCAACCAATTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTT 540
QY 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrLys 200
Db 541 TTATTAACCTTTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspLeu 220
Db 601 GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
Db 661 TTAGACTGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
Db 721 AATAAAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTAT 780
QY 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
Db 781 CATAGACACGACGAGTGGATGAACCAATTTGAACTTTGAACTTTGAACTTTGAACTTT 840
QY 281 LeuLeuGluIleValAspLeu 287
Db 841 TTATTAGAAATCGTTGATTTA 861

RESULT 2
US-09-925-637-63
; Sequence 63 Application US/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925, 637
```

Qy 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
Db 661 TTAGCTGCTATGTGTGATGAAGCAAGTTAGGTAAAGTGGCGCAGATCTTGAAGAA 720
Qy 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
Db 721 AATAAAGTACGTACGTGAGTTTATTAGGNAAGATGGCGCAGAGATAAATTGACTTAT 780
Qy 261 HisArgaspAlaAlaValaspGluLeuThrGlnIleaspGluGlnPheAsnThrLysHis 280
Db 781 CATAGACGCGCAGCGAGTGAAGTAACGCAAAATTGATGAACAATTCAATACAAAACAC 840
Qy 281 LeuLeuGluIleValaspLeu 287
Db 841 TTATTAGAAATCGTTGATTGA 861

RESULT 3

US-08-781-986A-155/c
; Sequence 155, Application US/08781986A
; Publication No. US20030054436A1

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248pp

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 155:

SEQUENCE CHARACTERISTICS:

LENGTH: 1893 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-155

Alignment Scores:

Pred. No.: 6.85e-161 Length: 1893
Score: 1453.00 Matches: 287
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-925-637-64 (1-287) x US-08-781-986A-155 (1-1893)

Qy 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValala 20
Db 1314 ATGACGAATCTACCGATGAATAAATAATAGATGAAGTCATAATGAATTATCGTGGC 1255
Qy 21 IleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAla 40

Db 1254 ATAATAAATCAGTAATGGATACAGCTAGAGAAAGTATGTTGTTATTCATTAATGCT 1195
Qy 41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
Db 1194 GGAGGTAACCCATCCGACAGTCTCTGTATTACTCCTTACTAGATTCACTAATACCCAG 1135
Qy 61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
Db 1134 TATGAGTTAGGTATGAGAGCGCAATTCAGTAAATGATTATACATATTCATTAT 1075
Qy 81 HisAspAspLeuProAlaMetAspAspAspTyrArgGlyLysLeuThrAsnHis 100
Db 1074 CATGATGACCTACCAGCGATGATTAATGATTATCGCAGGAGAAATTAACAATCAT 1015
Qy 101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
Db 1014 AAGTATATGCTGAGTGGCTGCGATATTAGCAGGTGATCTTTATTATTAACATAACAT 955
Qy 121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
Db 954 GAACCTTATTCAGTGTAGTATTAATCTGATGAAGTAAATAAAGTTCTACACGG 895
Qy 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
Db 894 CTGCAATAGCAAGTGGTCTATGTTGGAATGGTGGCGGTCAATGTTAGATATGCAAGC 835
Qy 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
Db 834 GAAGGCCAACCAATTGATCTTTGAACTTTGGAATGATACACAAAACAAACAGACGA 775
Qy 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrLys 200
Db 774 TTATTAACCTTTTGGGTTATGAGTGCAGCATATCGCTAATGTCGATGATACAACTAAA 715
Qy 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspLeu 220
Db 714 GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAAGATGATT 655
Qy 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
Db 654 TTAGCTGCTATGTTGATGAAGCAAGTTAGTAAAGTGGCGCAGCATCTTGAAAT 595
Qy 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
Db 594 AATAAAGTACGTACGTGAGTTTATTAGGAAAGATGCGCAGAGAGATAAATTCAGTTAT 535
Qy 261 HisArgaspAlaAlaValaspGluLeuThrGlnIleaspGluGlnPheAsnThrLysHis 280
Db 534 CATAGACGCGCAGCGATGATGAACCTAACGCAAAATTGATGAACAATTCAATACAAAAC 475
Qy 281 LeuLeuGluIleValaspLeu 287
Db 474 TTATTAGAAATCGTTGATTGA 454

RESULT 4

US-09-815-242-8485
; Sequence 8485, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21

;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 8485
;; LENGTH: 882
;; TYPE: DNA
;; ORGANISM: Staphylococcus aureus
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(882)
US-09-815-242-8485

Alignment Scores:
Pred. No.: 4,42e-160 Length: 882
Score: 1442.00 Matches: 285
Percent Similarity: 99.30% Conservative: 0
Best Local Similarity: 99.30% Mismatches: 2
Query Match: 99.24% Indels: 0
DB: 10 Gaps: 0

US-09-925-637-64 (1-287) x US-09-815-242-8485 (1-882)

QY 1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
Db 1 ATGACGAATCTACCGATGAATAAATTAAGATGAAGTCAATGAATATCGGTGG 60
QY 21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuThrValAsnAla 40
Db 61 ATAATAATACAGTAAATGATCTAGTACAGTAGAGAAAGATGTGTATTCATTAATGCT 120
QY 41 GlyGlySerGluThrProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 121 GGAGGTAAGCGCATCGGACGAGTCTGTATTACTCATTAGATTCACATAACCGAG 180
61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrThrSerLeuIle 80
181 TATGAGTTAGGTATGAGAGCGCAATGTCATAGAAATGATTCATATTCATTCATTTAT 240
81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
Db 241 CATGATGACCTACCGATGATATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 101 LysValTyrGlyGluThrProAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
Db 301 AAGATATATGTTGAGTGGATGCGATATGATGATGATGATGATGATGATGATGATGAT 360
QY 121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysLysLysValLeuGlnArg 140
Db 361 GAACCTATTTCAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
Db 421 CTGTCAATAGCAAGTGGTATGTTGAATGGTGGCGGTCAATGTTAGATATGCAAGC 480
QY 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
Db 481 GAAGGCCAACCAATTCATCTTGAACCTTTGGAATGATACACAAACAAACAGGAGCA 540
QY 181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200

Db 541 CTATTACATTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600
QY 201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
Db 601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTTCAGATTAAGATGATTTA 660
QY 221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAspLeuGluAsn 240
Db 661 TTAGACTGCTATGCTGATGAAGCGAGTTAGTAAAGTGGGACGCGATCTTGAATAA 720
QY 241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
Db 721 AATAAAGTATACATACGCTGAGTTTATTAGGAAAAGATGGCCAGAGATAAATTCAGCTAT 780
QY 261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
Db 781 CATAGAGACCGACGAGTGGATGAACTAATGCAAAATGATGAACAAATTCATAACAACAC 840
QY 281 LeuLeuGluIleValAspLeu 287
Db 841 TTATTAGAAATCGTTCATTTA 861
RESULT 5
US-09-815-242-4184
; Sequence 4184, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4184
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4184
Alignment Scores:
Pred. No.: 6,38e-157 Length: 864
Score: 1415.00 Matches: 280
Percent Similarity: 99.29% Conservative: 0
Best Local Similarity: 99.29% Mismatches: 2
Query Match: 99.38% Indels: 0
DB: 10 Gaps: 0
US-09-925-637-64 (1-287) x US-09-815-242-4184 (1-864)
QY 6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25

FILE REFERENCE: 10085.500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1015
LENGTH: 811
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-1015

Alignment Scores:
Pred. No.: 2,58e-65 Length: 811
Score: 635.00 Matches: 125
Percent Similarity: 73.11% Conservative: 49
Best Local Similarity: 52.52% Mismatches: 62
Query Match: 43.70% Indels: 2
Gaps: 10

US-09-925-637-64 (1-287) x US-09-974-300-1015 (1-811)

```
Qy 30 LeuGluSerMetLeuThrSerLeuAsnAlaGlyLysArgIleArgProValLeu 49
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 97 CTGAAGAAATCCATGCTGTAATGGAAGCGGGGGAACGGCTGAGACCGATATG 156
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 50 LeuLeuThrLeuAspSerLeuAsnThrGluThrGluLeuGlyMetLysSerAla 69
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 GTTCTTCGCGTTCATCGGTACGGAAGACGAGACCGGAAATTCAGTCGGATGC 216
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 70 AlaLeuGluMetIleHisThrSerLeuLeuHisAspLeuProAlaMetAspAsn 89
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 217 GCGTTGAATGATCCATGATGATTAATCCATGACGACCTCCCTTGTATGGATGAC 276
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 90 AspAspThrArgGlyGlyLeuThrAsnHisLysValThrGlyGluThrAlaIle 109
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 277 GATGACCTCAGAGAGGTAGCGGACAAACACCAAAATTTACGGGAGCGGCGCAT 336
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 110 LeuAlaGlyAspAlaLeuThrLysAlaPheGluLeuIleSerSerAsp----- 127
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 337 CTCGGGAGACGCTTATTAACGGAAGCTTCAAATGATCACTTCAATATGCTCTT 396
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 397 GATGATCCGCTGAAAGGCGCATCAGACTGGTCAACGAGCTGATTTCCGCGGCGGCC 456
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 457 GAAGGCGATGGTCGGCGGTGAGATTTAGATATGGAAGCGGAATCAAAGTCTGTATCT 516
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 168 ThrLeuThrGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 517 GATGAATCGCGCAATCCAGGAAGAAACCGCTAAACTCTCAGCTTCAGCGGTCAT 576
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 577 GCGGGAGCCATTCTTCAGATGATCTGAAAGAGAGATCGAAAGCTGCGCAATTCAGC 636
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuAspCysTyrGlyAspGlu 227
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 637 CATCATATCGCATCGGCTTCAGATCAGAGATGACATTTTGGACCTTGAAGGCTCGGAG 696
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 228 AlaLeuLeuGlyLysValGlySerAspLeuGluAsnLysSerThrTyrValSer 247
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 697 GATAAGATCGGCAACGAGATCGGATCCGAGCTTCAACGGAAGATTCGACGATTCGCTG 756
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAla 265
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 757 CTCCTTTCGCTTGAAGGCGGCCCAAAACCTTGATGAACATATTTGAAAGCA 810
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 8
US-09-815-242-6533
Sequence 6533, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6533
LENGTH: 882
TYPE: DNA
ORGANISM: Enterococcus faecalis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(882)
US-09-815-242-6533

Alignment Scores:
Pred. No.: 1,12e-60 Length: 882
Score: 596.00 Matches: 134
Percent Similarity: 65.29% Conservative: 56
Best Local Similarity: 46.05% Mismatches: 91
Query Match: 41.02% Indels: 10
DB: Gaps: 5

US-09-925-637-64 (1-287) x US-09-815-242-6533 (1-882)

```
Qy 1 MetThrAsnLeuProMetAsn-----LysLeuLeuAspGluValAsnLanGluLeuSer 18
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 ATGACCAATTTTAGTCAACAGCATTTACCGTTGGTTGAAAGTCAATGTCGATTTT 57
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeu 38
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 58 -----ATGCGAGAATATGCTGAAATGAGCGTTTGAAGAACCAATGCTCTACTCAATA 111
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 39 AsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsn 58
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 112 CACGCGAGTGGCAACGATTCGCGCGCTATTAGTGCTTAACACAGCTGGCGGCTTTTCAA 171
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 59 ThrGluTyrGluLeu---GlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyr 77
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 172 AAAGAGATGATGATGATTTACGAGCATATCAAGTGGCTCTTTAGAGATGATTCATACGCTAT 231
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 78 SerLeuIleHisAspAspLeuProAlaMetAspAsnAspTyrArgArgGlyLysLeu 97
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 232 TCATTAAATTCATGATGATTTACGAGCATGATGATGATGATGATGATGATGATGATGATGAT 291
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

QY 98 ThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThr 117
Db 292 ACCATCAAAAGTGTGGTGAAGCGACTGCCATTTTAGTGGGCGGCTTTATTAACA 351
QY 118 LysAlaPheGluLeuLeuSerSerAspArgLeuThrAspGluValLysIleLysVal 137
Db 352 GCGCATTTTACGTTTCTTGGCCAAATAGCTTAAGTGAA--AAAGTTTACTG 408
QY 138 LeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuAsp 157
Db 409 ATGCAACAACATGGCAAGCGAGCGGGAATCAAGCGCATGTTGCCGGCCAAATGGTGAT 468
QY 158 MetGlnSerGluGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLys 177
Db 469 ATTGAAGGAGAAAGTACAGCTTGACTTTTAGAAGAGTTAGCGGCTGTTTCATCAAAAGAAA 528
QY 178 ThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAsp 197
Db 529 ACCGGAGCGCTTAATGATTCGCTTAATTCGTTGGGGCGGTATTAGCCCAATCAACACAGA 588
QY 198 ThrThrLysGluHisLeuGluSerTyrThrHisLeuGlyMetMetPheGlnIleLys 217
Db 589 GAAGTCATTTGTTACTCACACAGTTTCCGCATCACTATGCTTGGCTTTTCAAAATCGT 648
QY 218 AspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAsp 237
Db 649 GATGACTTATTAGATGCCACAGTACAGACAGACATTTAGTAAAGAAAGTAGCGCGAGAT 708
QY 238 LeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLys 257
Db 709 GAAGGTTGATTAAGTACGTATCCAGCTCTTCTAGGGATGTTGGCGCGCAAGATCGG 768
QY 258 LeuThrTyr-----HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGlu 274
Db 769 CTAACACATCAATTAGCAGAGAGCGCTGTTTAGAAAAAATCAAGCAAAATGTTCCC 828
QY 275 GlnPheAnThrLysHisLeuLeuLulVal 285
Db 829 AACTTTTCAGAGAACAATTAGCAAACTTGTTA 861

RESULT 9

US-09-070-927A-55
; Sequence 55, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Steven Barash
; Patrick J. Dillon
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 7528 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-070-927A-55
Alignment Scores:
Pred. No.: 2,528-58 Length: 7528
Score: 587.50 Matches: 130
Percent Similarity: 67.68% Conservative: 48
Best Local Similarity: 49.43% Mismatches: 78
Query Match: 40.43% Indels: 7
DB: 10 Gaps: 4
US-09-925-637-64 (1-287) x US-09-070-927A-55 (1-7528)
QY 1 MetThrAsnLeuPrometAsn-----LysLeuIleAspGluValAsnAsnGluLeuSer 18
Db 6642 ATGACGAATTTTAGTCAACAGCATTTACCGTTGGTTGTTGAAAGATCATGTCGATTTT--- 6698
QY 19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeu 38
Db 6699 -----ATCGCAGAAATATCTGAAATGACCGTTTGAAGAGCAATGCTCTACTCAATA 6752
QY 39 AsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 58
Db 6753 CACGCGAGTGGCAACAGATAGCCCGCTATTAGTGTTAACACAGTAGTCCGCCCTTTTCAA 6812
QY 59 ThrGluTyrGluLeu---GlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyr 77
Db 6813 AAAGAGATGGAACACACAGACTATCAAGTGGCTCTCTTTAGAGATGATTCATACGTAT 6872
QY 78 SerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeu 97
Db 6873 TCATTAAATTCATGATGATTTACCAAGATGAGATGATGATTTACGTCGTCGCAACCA 6932
QY 98 ThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuThr 117
Db 6933 ACCAATCATAAAGTGTGGTGAAGCGACTGCCATTTTAGCAGGGGACGGCTTTATTAACA 6992
QY 118 LysAlaPheGluLeuLeuSerSerAspArgLeuThrAspGluValLysIleLysVal 137
Db 6993 GGTGCATTTTTCAGTTCTTCTTTGAGTCAATTAGGCTTAAGTGAA---AAAGTTTACG 7049
QY 138 LeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeuAsp 157
Db 7050 ATGCAACAACATGGCAAGCGAGCAATCAAGCATGCTTCCCGCAATGGGTGAT 7109
QY 158 MetGlnSerGluGlnProIleAspLeuGluThrLeuGluMetIleHisThrLys 177
Db 7110 ATTGAAGGAGAAAGATCAGCTTGACTTTAGAGAGTAGTACGCGCTTCTCAAGAGAAA 7169
QY 178 ThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAsp 197
Db 7170 ACCGGAGCGCTAAATGAAATTCGCTTAATTTGGGGGGCGGTATTAGCCCAATCAACAGAA 7229
QY 198 ThrThrLysGluHisLeuGluSerTyrThrHisLeuGlyMetMetPheGlnIleLys 217
Db 7230 GAAGTCATTTGGTTTACTCACACAGTTTCCGCATCACTATGCTTGGCTTTTCAAAATCGT 7289
QY 218 AspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySerAsp 237
Db 7290 GATGACTTATTAGATGCCCAAGTAGCAGAGCAGATTTAGGTAAAGAGTAGCGCGAGAT 7349

Db 193 GTACCTACTGCTTGGAGATGATTCATACAGGAGGCTTGATTCACGATGACCTTCCTGCT 252
Qy 87 MetAspAsnAspTyrArgGlyLysLeuThrAsnHisLysValTyrGlyLysTrp 106
Db 253 ATGGATGATGACGATTTACGAAGGCGGTAAACCAATCACAAGAAATTCGGTGAAGCT 312
Qy 107 ThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuLeuSerSerAsp 126
Db 313 ATGGCCATTTGGCTGGAGATGCTTATCTTACACCATATGCTTGTGATGGCAGCA 372
Qy 127 AspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGly 146
Db 373 GAT---TTGCCAGTCAGATTAAGTGGAGCTTGATTCACCACTATCCCTGCTTCAGT 429
Qy 147 HisValGlyMetValGlyGlnMetLeuAspMetGlnSerGlyGlnProIleAsp 166
Db 430 AGTCGGTATGTTGGCAGGCAAGTTTGGATATGGAGGCGGAACACACGACCTTCTCT 489
Qy 167 LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal 186
Db 490 TTGGAAGAACTTCAGATTTATCATGCAATTAAGACTGGGAAGTTACTAGCTATCCCTTC 549
Qy 187 MetSerAlaAlaAspIleAlaAsnValAspAspThrLysGluHisLeuGluSerTyr 206
Db 550 CAAGCGCAGCTATTATAGCTGATTCACCTGAATGCAAGTGAAGTGAAGTGAAGTGA 609
Qy 207 SerTyrHisLeuGlyMetMetPheGlnIleLysAspLeuLeuAspCysTyrGlyAsp 226
Db 610 GGTGAATGATTTGGCTGGAGCTTGCTTCAAGTCAGAGATGATGATGATGATGATGAT 669
Qy 227 GluAlaLysLeuGlyLysValGlySerAspLeuGluAsnAsnLysSerThrTyrVal 246
Db 670 TTTGAGGAATTCGGTAAAGACACCTCAAAAGGATGTCGAGGAGGAGGAGGAGGAGG 729
Qy 247 SerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaVal 266
Db 730 GCCTTGTGGCTTGGAGAGTCCATTCCTTTTGTAAACCAACCTGGATCAAGCTAAT 789

RESULT 12

US-09-815-242-9175

Sequence 9175, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 9175
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(876)
US-09-815-242-9175
Alignment Scores:
Pred. No.: 6,32e-55 Length: 876
Score: 547.00 Matches: 120
Percent Similarity: 64.89% Conservative: 50
Best Local Similarity: 45.80% Mismatches: 88
Query Match: 37.65% Indels: 4
DB: 10 Gaps: 3

US-09-925-637-64 (1-287) x US-09-815-242-9175 (1-876)

Qy 28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyLysArgIleArgPro 47
Db 73 TCTAGTTTACGGGAGTCTGTTCTCTATTCATTCATGCTGCTGCAACGATTCGGCT 132
Qy 48 ValLeuLeuLeuLeuThrLysSerLeuAsnThrGluTyrGlu---LeuGlyMetLys 66
Db 133 TTTCTCTGTTAGAGTTCTGGAAGCTTGCAGGTTGCCATCAACCTGCACGCGAG 192
Qy 67 SerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAla 86
Db 193 GTACCTACTGCTTGGAGATGATTCATACAGGAGCTTGATTCAGATGACCTTCCTCT 252
Qy 87 MetAspAsnAspTyrArgGlyLysLeuThrAsnHisLysValTyrGlyLysTrp 106
Db 253 ATGGATGATGACGATTTATCGAAGGCGGTAAACCAATCACAAGAAATTCGGTGAAGCT 312
Qy 107 ThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuLeuSerSerAsp 126
Db 313 ATGCCATTTGGCTGGAGATGCTTATTCATACACCATATGCTTGTGATGGCAGCA 372
Qy 127 AspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGly 146
Db 373 GAT---TTGCCAGTCAGATTAAGTGGAGCTTGATTCACCACTATCCCTGCTTCAGT 429
Qy 147 HisValGlyMetValGlyGlnMetLeuAspMetGlnSerGlyGlnProIleAsp 166
Db 430 AGTCGGTATGTTGGCAGGCAAGTTTGGATATGGAGGCGGAACACACGACCTTCTCT 489
Qy 167 LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal 186
Db 490 CTGGAAGAACTTCAGACTATTTCATGCCCAATTAAGACTGGGAAGTTACTAGCTATCCCTTC 549
Qy 187 MetSerAlaAlaAspIleAlaAsnValAspAspThrLysGluHisLeuGluSerTyr 206
Db 550 CAAGCGCAGCTATTATAGCTGATTCACCTGAATGCAAGTGAAGTGAAGTGAAGTGA 609
Qy 207 SerTyrHisLeuGlyMetMetPheGlnIleLysAspLeuLeuAspCysTyrGlyAsp 226
Db 610 GGTGAATGATTTGGCTGGAGCTTGCTTCAAGTCAGAGATGATGATGATGATGATGAT 669
Qy 227 GluAlaLysLeuGlyLysValGlySerAspLeuGluAsnAsnLysSerThrTyrVal 246
Db 670 TTTGAGGAATTCGGTAAAGACACCTCAAAAGGATGTCGAGGAGGAGGAGGAGGAGG 729
Qy 247 SerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaVal 266
Db 730 GCCTTGTGGCTTGGAGAGTCCATTCCTTTTGTAAACCAACCTGGATCAAGCTAAT 789

Qy 63 LeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAsp 82
Db 196 GTGCTGGATGCGCGGCTTGGCGGTGAGAAATTCATCCATGTGTTATCGCTGATTCACGAC 255
Qy 83 AspLeuProAlaMetAspAsnAspTyrArgGlyLysLeuThrAsnHisLysVal 102
Db 256 GATCTGCGCGCATGGACACGATGATCTGCGCGCGGCAACGACCTGTCACAGGCT 315
Qy 103 TyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeu 122
Db 316 TACGACGAGCGACCGCATTTTGGCGCGGACGACCTGCGAGGCGCTTGAAGTT 375
Qy 123 IleSerSerAspArgLeuThrAspGlu-----ValLysIleLysValLeuGlnArg 140
Db 376 CTGCGCAACGACCGCGCATCAGCTGATGCGCGCGCTCGCTGCAAAATGATCAGCGCT 435
Qy 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
Db 436 TTGACCGCGCGCAGCGCTCTCAAGGATGTTGGCGGTCAAGCCATGATCTCGGCTCC 495
Qy 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
Db 496 GTCGCGCGCAAAATGACGCTCGCGGAACCTCGAAACATGATATCCACAGACTGGCGCC 555
Qy 181 LeuLeuThrPheAlaVal---MetSerAlaAlaAspIleAlaAsnValAspAspThr 199
Db 556 CTGATCGCGCGCAGCGCTCAATCTGGCGCATATTCGCGCGCTGCGCTGCGCTGCGCTC 615
Qy 200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
Db 616 GCCAAGAACTGGATCATCTGCAAGGATGTTGGCGGTCAAGCCATGATCTCGGCTCC 495
Qy 220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
Db 676 ATTCGACATCGAAGCGCACCGCGACACTCGGCAAGACTCAGGCGAAGGACATCGAT 735
Qy 240 AsnAsnLysSerThrTyrValSerLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
Db 736 AACGACAAACCGACCTACCTCGCGCTATTTGGCGCATGCTGGCGCGCAACAAAGCCGAG 795
Qy 260 TyrHisArgAspAlaAlaValAspLysLeuThr 270
Db 796 GAATTCGACGAAACGACGAGTCAAGCTTAACG 828

RESULT 15

us-09-934-868-71
; Sequence 71, Application US/09934868

Patent No. US20020137190A1

GENERAL INFORMATION:

APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97

SEQ ID NO 71

LENGTH: 891

TYPE: DNA

ORGANISM: Methylomonas 16a

FEATURE:

OTHER INFORMATION: ORF7 ISPa

US-09-934-868-71

Alignment Scores:

Pred. No.: 2,48e-51 Length: 891
Score: 516.50 Matches: 113
Percent Similarity: 63.75% Conservative: 47

Best Local Similarity: 45.02% Mismatches: 88

Query Match: 35.55% Indels: 3

DB: 10 Gaps: 2

US-09-925-637-64 (1-287) x US-09-934-868-71 (1-891)

Qy 23 LysSerValMetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAlaGlyGly 42
Db 76 GAAACATATCTGCGCAACAACTTGCATCAGGCCATGCGCTATTCGTTATGAACGCGCGC 135
Qy 43 LysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGlu 62
Db 136 AACGACCGCGCGCTTGTGACTTATCGGACCGCTGAGGCTTGGGCTTGGCGGAAAC 195
Qy 63 LeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAsp 82
Db 196 GTGCTGGATGCGCGGCTTGGCGGTGAGAAATTCATCCATGTGTTATCGCTGATTCACGAC 255

Qy 83 AspLeuProAlaMetAspAsnAspTyrArgGlyLysLeuThrAsnHisLysVal 102
Db 256 GATCTGCGCGCATGGACACGATGATCTGCGCGCGGCAACGACCTGTCACAGGCT 315
Qy 103 TyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeu 122
Db 316 TACGACGAGCGCACCGCATTTTGGCGCGGACGACCTGCGAGGCGCTTGAAGTT 375
Qy 123 IleSerSerAspArgLeuThrAspGlu-----ValLysIleLysValLeuGlnArg 140
Db 376 CTGCGCAACGACCGCGCATCAGCTGATGCGCGCGCTCGCTGCAAAATGATCAGCGCT 435

Qy 141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
Db 436 TTGACCGCGCGCAGCGCTCTCAAGGATGTTGGCGGTCAAGCCATGATCTCGGCTCC 495
Qy 161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
Db 496 GTCGCGCGCAAAATGACGCTCGCGGAACCTCGAAACATGATATCCACAGACTGGCGCC 555
Qy 181 LeuLeuThrPheAlaVal---MetSerAlaAlaAspIleAlaAsnValAspAspThr 199
Db 556 CTGATCGCGCGCAGCGCTCAATCTGGCGCATATTCGCGCGCTGCGCTGCGCTGCGCTC 615

Qy 200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
Db 616 GCCAAGAACTGGATCATCTGCAAGGATGTTGGCGGTCAAGCCATGATCTCGGCTCC 495
Qy 220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
Db 676 ATTCGACATCGAAGCGCACCGCGACACTCGGCAAGACTCAGGCGAAGGACATCGAT 735
Qy 240 AsnAsnLysSerThrTyrValSerLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
Db 736 AACGACAAACCGACCTACCTCGCGCTATTTGGCGCATGCTGGCGCGCAACAAAGCCGAG 795

Qy 260 TyrHisArgAspAlaAlaValAspLysLeuThr 270
Db 796 GAATTCGACGAAACGACGAGTCAAGCTTAACG 828

Search completed: May 30, 2003, 22:36:09

Job time : 192 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 20:17:41 ; Search time 1557 Seconds
(without alignments)
2985.295 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMMKLIDEVNNELSA.....ELTQIDRQFNKHLLEIVDL 287

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Sorted: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09925637/runat_23052003.174818.23437/app.query.fasta_1.455
-DB=EST -QFMT=fastap -SUFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637 -ECGN_1_1906.erunat_23052003.174818.23437 -NCPU=6 -ICPU=3
-NO_MMALP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

- Database :
- 1: em_estba:**
 - 2: em_esthum:**
 - 3: em_estin:**
 - 4: em_estmu:**
 - 5: em_estov:**
 - 6: em_estpl:**
 - 7: em_estro:**
 - 8: em_htc:**
 - 9: gb_est1:**
 - 10: gb_est2:**
 - 11: gb_htc:**
 - 12: gb_est3:**
 - 13: gb_est4:**
 - 14: gb_est5:**
 - 15: em_estfun:**
 - 16: em_estom:**
 - 17: gb_gss:**
 - 18: em_gss_hum:**
 - 19: em_gss_inv:**
 - 20: em_gss_pln:**
 - 21: em_gss_vrt:**
 - 22: em_gss_fun:**
 - 23: em_gss_mam:**
 - 24: em_gss_mus:**
 - 25: em_gss_othr:**
 - 26: em_gss_pro:**
 - 27: em_gss_rtd:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	533.5	36.7	4484	17	BH770981	BH770981 LLMGTaq70
2	436.5	30.0	682	14	BQ856404	BQ856404 QGB3904.Y
3	431	29.7	685	13	BM449136	BM449136 DS032E03
4	418	28.8	703	17	BH429531	BH429531 BOHLB93TR
5	415	28.6	709	17	BH502634	BH502634 BOHJK46TR
6	378.5	26.0	797	17	BH466685	BH466685 BOGOW16TR
7	378.5	26.0	863	17	BH584613	BH584613 BOGOW60TR
8	356.5	24.5	475	17	BH439236	BH439236 BOHAC18TF
9	343	23.6	837	12	BF261534	BF261534 HV_CERA000
c 10	339.5	23.4	770	17	BH250143	BH250143 BOGAG35TF
c 11	336.5	23.2	442	9	AJ302129	AJ302129 AJ302129
c 12	325.5	22.4	808	14	BQ506777	BQ506777 EST614192
c 13	314	21.6	596	17	AQ963747	AQ963747 LERGP06TF
c 14	314	21.6	838	17	BH657017	BH657017 BOMBM36TF
15	299	20.6	666	13	BI717621	BI717621 1031021A0
16	299	20.6	675	13	BI723806	BI723806 1031067H0
17	299	20.6	701	10	BE238154	BE238154 894040507
18	298.5	20.5	657	9	AJ502290	AJ502290 AJ502290
19	297.5	20.5	425	10	BE323612	BE323612 NF006A09P
20	297.5	20.5	558	9	AL749888	AL749888 AL749888
21	297.5	20.5	828	17	BH501300	BH501300 BOHAP61TF
22	297	20.4	384	10	BE521431	BE521431 M19E5STM
23	296	20.4	669	17	BH731669	BH731669 BOHYD34TF
24	292	20.1	668	13	BI726915	BI726915 1031088E0
25	291.5	20.1	571	10	BE432698	BE432698 EST399227
26	291.5	20.1	603	9	AI855654	AI855654 sc32d02.Y
27	288.5	19.9	559	14	BQ864712	BQ864712 QGE27H09
28	288.5	19.9	680	14	BQ983434	BQ983434 QGE19B22
29	288.5	19.9	712	14	BH007746	BH007746 QGH5F18.Y
30	288.5	19.9	720	14	BQ984560	BQ984560 QGE2h02.Y
c 31	288	19.8	498	17	AQ963748	AQ963748 LERGP06TR
c 32	286	19.7	576	13	BI944945	BI944945 sa124G02
c 33	284.5	19.6	771	17	BH479446	BH479446 BOGTW38TF
34	283.5	19.5	468	12	BF519441	BF519441 EST456904
35	282.5	19.4	675	14	BQ983788	BQ983788 QGE1c08.Y
36	280	19.3	561	12	BF054034	BF054034 EST439264
37	280	19.3	636	9	AU238772	AU238772 AU238772
38	279	19.2	402	10	AW719932	AW719932 LJNEST12b
39	278	19.1	698	13	BI946026	BI946026 st91f10.Y
40	274	18.9	722	14	BQ916912	BQ916912 QHB19009
41	272	18.7	355	9	AI780894	AI780894 EST261773
42	271	18.7	735	14	BQ847849	BQ847849 QGA5117.Y
43	269	18.5	628	14	BQ982915	BQ982915 QGE17L12
44	269	18.5	716	14	BQ984280	BQ984280 QGR21D23
45	269	18.5	745	14	BQ991706	BQ991706 QGR23J13

ALIGNMENTS

RESULT 1
BH770981
LOCUS BH770981 4884 bp DNA linear GSS 01-MAY-2002
DEFINITION LLMGTaq706 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, DNA sequence.
ACCESSION BH770981
VERSION BH770981.1 GI:20373938
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 4484)
AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis


```
Pred. No.: 2.39e-45 Length: 682
Score: 436.50 Matches: 97
Percent Similarity: 62.39% Conservative: 44
Best Local Similarity: 42.92% Mismatches: 76
Query Match: 30.04% Indels: 9
DB: 14 Gaps: 3

US-09-925-637-64 (1-287) x BQ856404 (1-682)

Qy 16 GluLeuSerValAlaLeuAsnLysSerValMetAspThr----- 28
Db 5 GAGAAAGCTCTCGGGTTAAGAGCGCTTGATGATCCATTTCCATTAAGAGCCACCC 64
Qy 29 GlnLeuGluSerMetLeuThrSerLeuAsnAlaGlyGlyLysArgIleArgProVal 48
Db 65 ACAATCATCAAGCTATGCTTACTCTCTCGCGGTGGGAAGCTGTGAGGCCCAT 124
Qy 49 LeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluLeuGlyMetLysSerAla 68
Db 125 CTCTGCATCCCGCTTCGAGCTCGTCGGAGGTGATGAATCCACCGCCATGCCAGCGCC 184
Qy 69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAsp 88
Db 185 TGGCGCTTGAATGATACACCATCTGCTGATTCACGACGACCTTCGCTGCATGGAT 244
Qy 89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyLeuThrAla 108
Db 245 AAGCAGCATTTTCGCGGGGGAACCAACCAACCAAGGTTTACGCGGAAGAGCTGGCG 304
Qy 109 IleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAsp 127
Db 305 GTCTTCGCGCGGACTCTCTCTCGCTTCACACATCGCAAGCGCCACCGCTC 364
Qy 128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
Db 365 GCGCGCTCTCTGAACCGTTTGGCGCGCTCGGAGAACTGGCGAAGTGGGACT 424
Qy 148 ValGlyMetValGlyGlnMetLeuAspMetGlnSerGluGly---GlnProIleAsp 166
Db 425 GAAGGGTGTGCGGGTCAAGTGGTGACATAGCTTCAACAGGTGCAAGAGCTCGGA 484
Qy 167 LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal 186
Db 485 CTGGACAGCTGTGATTCATCCACATCCACAGCGCGCTATTGGAGAGCGCTC 544
Qy 187 MetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyr 206
Db 545 GTATTGGAGCAATATTAGTGGAGGAGGACACACCAAGTGGAAAGCTGAGAAAGTTC 604
Qy 207 SerTyrHisLeuGlyMetMetPheGlnIleLysAspLeuLeuAspCysTyrGlyAsp 226
Db 605 GCGAGATGCATTGGATTATTGTTTTCAGGTGGTTGACGACATATTAGACGTCACCAAGTCG 664
Qy 227 GluAlaLysLeuGlyLys 232
Db 665 TCGGAGGAATGGGAAA 682

RESULT 3
BM449136
LOCUS
DEFINITION
685 bp mRNA linear EST 01-APR-2002
DSAO32E03_60893 An expressed sequence tag database for the
halotolerant green alga, Dunaliella salina Dunaliella salina cDNA
clone DSAO32E03 5, mRNA sequence.
ACCESSION
BM449136
VERSION
BM449136.1 GI:19854708
KEYWORDS
EST.
SOURCE
Dunaliella salina.
ORGANISM
Dunaliella salina
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Dunaliellaceae; Dunaliella.
REFERENCE
1 (bases 1 to 685)
Cushman,J.C.
AUTHORS
An expressed sequence tag database for the halotolerant green alga,
TITLE
```

JOURNAL
COMMENT

Dunaliella salina
Unpublished (2002)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers

FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 032 row: E column: 03
Seq primer: T3 20mer
High quality sequence stop: 685.
Location/Qualifiers

FEATURES
source

1..685
/organism="Dunaliella salina"
/db_xref="taxon:3046"
/clone="DSAO32E03"
/clone_lib="An expressed sequence tag database for the
halotolerant green alga, Dunaliella salina"
/tissue_type="Cells, which was adapted in 2.5M NaCl via a
incremental series from 1.7 to 2.0 to 2.25 to 2.5 M NaCl,
were exposed to 3.4 M NaCl for 5 hours"
/cell_type="Green"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Library construction was performed
according to Stratagene's recommended protocol for the
Lambda UniZapXR vector and cDNA synthesis kit."
BASE COUNT 153 a 182 c 219 g 131 t
ORIGIN

Alignment Scores:

Pred. No.: 1.23e-44 Length: 685
Score: 431.00 Matches: 92
Percent Similarity: 63.58% Conservative: 36
Best Local Similarity: 45.77% Mismatches: 61
Query Match: 29.66% Indels: 12
DB: 13 Gaps: 3

US-09-925-637-64 (1-287) x BM449136 (1-685)

Qy 78 SerLeuIleHisAspAspLeuProAlaMetAspAsnAspTyrArgArgGlyLysLeu 97
Db 7 AGCCTGATCCATGATGACTTCCCTCCATGACACGATGACTTCCGAGAGCGGCC 66
Qy 98 ThrAsnHisLysValTyrGlyGluThrPheAlaIleLeuAlaGlyAspAlaLeuThr 117
Db 67 ACCAACCAAGGTGACGAGAGGACATTCCTGCTGCTGAGATGCCCTCTCTCT 126
Qy 118 LysAlaPheGluLeuIleSer---SerAspAspArgLeuThrAspGluValLysLys 136
Db 127 TTCGCTTGGACATGTTCGCGCGGACCACTGCCCTCGCTGACGCTGCTGGT 186
Qy 137 ValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlnMetLeu 156
Db 187 GTGATTCTGGAGCTGGGCAAGGCTGTGTCAGATGGCTGCACAGGAGGACAGT 246
Qy 157 AspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLys 176
Db 247 GACATTAGACGGAAGAGGAGGAGGTGGGCTGGAGGTGCTGACATAGATCCAT 306
Qy 177 LysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnVal 196
Db 307 AAGACAGCAGCCCTGCTGGAGCATCCGTGTGTGCTGCGCTCTCTGGGCGG 366
Qy 197 AspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGln 216
Db 367 GACGTGACAGTGGAGAGCTGCGCAAGTACGCGCAACATTTGGCTCGCTCCAG 426
Qy 217 LysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGly 236

```

Db      427 GTTGACGATATCTCGATTGACCCAGACACAGGAGATGCTGGGCAAGAGCGGAGCAAG 486
QY      237 AspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGly-----LysAsp 252
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      487 GACATTGATGTGAACAAGACACAGGTACCCCAAACTGCTGGCGTGGAGAGAGTAAGCAG 546
QY      253 GlyAlaGluAspLysLeuThrTyrHisArgAlaValAspGluLeuThrGlnIle 272
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      547 GCTGCAGAGGAGCCTGATT-----GCTGAGGCCATCCAGCAGCTG 585
QY      273 Asp 273
      |||
Db      586 GAT 588

RESULT 4
LOCUS   BH429531
DEFINITION BOHLB93TR BOHL Brassica oleracea genomic clone BOHLB93, DNA
       703 bp DNA linear GSS 12-DEC-2001
       sequence.
       BH429531
       BH429531
       BH429531.1 GI:17615252
       GSS.
       SOURCE Brassica oleracea.
       ORGANISM Brassica oleracea
       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
       Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
       Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
       1 (bases 1 to 703)
       Whole genome shotgun sequencing of Brassica oleracea
       Unpublished (2001)
       Other_GSSs: BOHLB93TF
       Contact: Chris Town
       TIGR
       9712 Medical Center Drive, Rockville, MD 20850, USA.
       Tel: 301-838-3523
       Fax: 301-838-0208
       Email: cdtown@tigr.org
       DNA is from a doubled haploid provided by Tom Osborn.
       Seq primer: TR
       Class: sheared ends.
       Location/Qualifiers
       1..703
       /organism="Brassica oleracea"
       /strain="TO1000DH3"
       /db_xref="taxon:3712"
       /clone_lib="BOHLB93"
       /note="Vector: PHOS1; Site_1: BstXI; 2-3 kb sheared
       genomic DNA inserted into PHOS1 using BstXI linkers."
BASE COUNT 180 a 143 c 216 g 164 t
ORIGIN
Alignment Scores:
Pred. No.: 5,99e-43 Length: 703
Score: 418.00 Matches: 95
Percent Similarity: 60.09% Conservative: 39
Best Local Similarity: 42.60% Mismatches: 85
Query Match: 28.77% Indels: 4
DB: 17 Gaps: 2

US-09-925-637-64 (1-287) x BH429531 (1-703)

QY      29 GlnLeuGluSerMetLeuTyrSerLeuAsnAlaGlyLysArgIleArgProval 48
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      37 AAGATCCAGGAGCCATGCTTACCGGTTCTTCGACGAGGAGAAACCGCTTAGGCCAATA 96
QY      49 LeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAla 68
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      97 CTATGCTTGTCTTTCGCGAGTAGTAGGAGCGCAAGAAACCGCGCGAGCGGCT 156
QY      69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAsp 88
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

```

```

Db      157 TGTGCGGTGAGATGATACACACCATGTCTCTTAATCAAGACGACTGCTTGTATGGAC 216
QY      89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluThrPheAla 108
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      217 AATGACGATCTGCTGCTGGGAAAGCCTACGACCCACAAAGTCTACGCGGAAGAGTCGCT 276
QY      109 IleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuLeuSerSerAspAspArg 128
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      277 ATTCTCTCCGGAGGAGCTCTCTTGTCTTCTGCTTCGAGCAGCATGACGACGGCGGAG--- 333
QY      129 LeuThrAspGluValLysLysValLeuGlnArgLeuSerIleAlaSerGlyHisVal 148
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      334 GTATCTCGGAGAAATGTTTGGCGGCTGAGGAACTGCTAGGTCCATTGGAACATAGA 393
QY      149 GlyMetValGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp----- 166
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      394 GGGTTAGTCCGGGACACAGCTAAGGACATAAGTAGTGAAGTTTGGACTTAGACGAGGTA 453
QY      167 ---LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAla 185
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      454 GGACTAGACATTTAGAGTTTATACACGTACACAAACCGCGGCTTCTTTTGGAAACTGCT 513
QY      186 ValMetSerAlaAlaAspIleAlaAsnValAspThrThrLysGluHisLeuGluSer 205
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      514 GCGGTTCTTGGAGCCATTATTGCTGGGTCTTAATGAAGAGCTTGAGAGAGTTAGAAAG 573
QY      206 TyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGly 225
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      574 TTGCGCAAGTGCATTCGGTTGTTTCAAGTGGTGGATGATATTTTGGACGACACCAAG 633
QY      226 AspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyr 245
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      634 TCGTCGAGGAACCTGGGAAACCCGCGGAAAGATCAGCTCTCGGAAAGCTGACTTAT 693
QY      246 ValSerLeu 248
      |||
Db      694 CCCAAGCTG 702

RESULT 5
BH502634
LOCUS   BOHJK46TR BOHJ Brassica oleracea genomic clone BOHJK46, DNA
DEFINITION
       709 bp DNA linear GSS 13-DEC-2001
       sequence.
       BH502634
       BH502634.1 GI:17710731
       GSS.
       SOURCE Brassica oleracea.
       ORGANISM Brassica oleracea
       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
       Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
       Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
       1 (bases 1 to 709)
       Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
       Whole genome shotgun sequencing of Brassica oleracea
       Unpublished (2001)
       Other_GSSs: BOHJK46TF
       Contact: Chris Town
       TIGR
       9712 Medical Center Drive, Rockville, MD 20850, USA.
       Tel: 301-838-3523
       Fax: 301-838-0208
       Email: cdtown@tigr.org
       DNA is from a doubled haploid provided by Tom Osborn.
       Seq primer: TR
       Class: sheared ends.
       Location/Qualifiers
       1..709
       /organism="Brassica oleracea"
       /strain="TO1000DH3"
       /db_xref="taxon:3712"
       /clone_lib="BOHJK46"
       /note="Vector: PHOS1; Site_1: BstXI; 2-3 kb sheared
       genomic DNA inserted into PHOS1 using BstXI linkers."
BASE COUNT 180 a 143 c 216 g 164 t
ORIGIN
Alignment Scores:
Pred. No.: 5,99e-43 Length: 703
Score: 418.00 Matches: 95
Percent Similarity: 60.09% Conservative: 39
Best Local Similarity: 42.60% Mismatches: 85
Query Match: 28.77% Indels: 4
DB: 17 Gaps: 2

US-09-925-637-64 (1-287) x BH429531 (1-703)

QY      29 GlnLeuGluSerMetLeuTyrSerLeuAsnAlaGlyLysArgIleArgProval 48
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db      37 AAGATCCAGGAGCCATGCTTACCGGTTCTTCGACGAGGAGAAACCGCTTAGGCCAATA 96
QY      49 LeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAla 68
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      97 CTATGCTTGTCTTTCGCGAGTAGTAGGAGCGCAAGAAACCGCGCGAGCGGCT 156
QY      69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuProAlaMetAsp 88
      |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

```


TITLE Whole genome shotgun sequencing of Brassica oleracea
 COMMENT Unpublished (2001)
 Other_GSSs: BOGWJ16TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends

FEATURES
 source
 1..797
 /organism="Brassica oleracea"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BOGWJ16"
 /clone_lib="BOGW"
 /note="Vector: PHOS1; Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 167 a 139 c 248 g 155 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.48e-42 Length: 709
 Score: 415.00 Matches: 91
 Percent Similarity: 60.45% Conservatve: 42
 Best Local Similarity: 41.36% Mismatches: 83
 Query Match: 28.56% Indels: 4
 DB: 17 Gaps: 2

US-09-925-637-64 (1-287) x BH502634 (1-709)
 Qy 40 AlaGlyGlyLysArgIleAraProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThr 59
 Db 3 GCCGGCGGCAACGCGTCAGACCAATACTCTCGCTCGCGGCTCGGAGTAGTCGGAGG 62
 Qy 60 GluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeu 79
 Db 63 GAAGAGCGTTCGGCGGTTCCGGCGGCTCGCGGTGAGATGATTCACAGATGTCAC 122
 Qy 80 IleHisAspLeuProAlaMetAspAspAspTyrArgArgGlyLysLeuThrAsn 99
 Db 123 ATCAAGACGCTCTCTTCATGGACACGACGACCTCGCGGAGTAAAGCCACGACG 182
 Qy 100 HisLysValTyrGlyGluThrPThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAla 119
 Db 183 CACAAAGCTCTCGGAGAAAGCGTCGCGATCTCTCCGGCGCGGCTCTAGCTCTGCC 242
 Qy 120 PheGluLeuLeuSerSerAspArgLeuThrAspGluValLysIleLysValLeuGln 139
 Db 243 TTCAGCATTTGCGGAAGCTGAC--GTGTGCGCTGAGAGATGGTTAGACGGTTAAA 299
 Qy 140 ArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGln 159
 Db 300 GAATGCGGAGTCTATAGGACGAAAGGCTCGTGGCGGACAAAGCGATGGATTGAGC 359
 Qy 160 SerGluGly-----GlnProIleAspLeuGluThrLeuGluMetIleHisLysThr 176
 Db 360 AGCGAAGTTTGGATCAAAACGACGCTCGGTTGAAGGAGCTTCAAGTTTATTCAGTTCAT 419
 Qy 177 LysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAsp 196
 Db 420 ARAACCGTTCGCTGTGAGGCTTCGCGGTTATTCGAGCGGTATTCGAGGTGGTTCG 479
 Qy 197 AspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIle 216
 Db 480 GAGGAAGGTTGAGAGGTGAGGAGCTTCGCGAGGTGTATGGGTGTGTTTCAGGTG 539
 Qy 217 LysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysValGlySer 236
 Db 540 GTTGATGATATTTGGATGGACTTAAGTCGTCGGAGGAGCTGGGGAAGACCGCGGAAA 599
 Qy 237 AspLeuGluAanAsnLysSerThrTyrValSerLeuGlyLysAspGlyAlaGluAsp 256
 Db 600 GATCAGGTGCGGGAAGCTGACGTATCCGAAGGTGTGGACTTGAGGAAGTAAAGAG 659

RESULT 6
 BH466685 797 bp DNA linear GSS 13-DEC-2001
 LOCUS BOGWJ16TR BOGW Brassica oleracea genomic clone BOGWJ16, DNA
 DEFINITION sequence.
 ACCSSION BH466685
 VERSION BH466685.1 GI:17667767
 KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 797)
 AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.

TITLE Whole genome shotgun sequencing of Brassica oleracea
 COMMENT Unpublished (2001)
 Other_GSSs: BOGWJ16TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends

FEATURES
 source
 1..797
 /organism="Brassica oleracea"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BOGWJ16"
 /clone_lib="BOGW"
 /note="Vector: PHOS1; Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 187 a 190 c 237 g 183 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.79e-38 Length: 797
 Score: 378.50 Matches: 92
 Percent Similarity: 54.04% Conservatve: 35
 Best Local Similarity: 39.15% Mismatches: 103
 Query Match: 26.05% Indels: 5
 DB: 17 Gaps: 1

US-09-925-637-64 (1-287) x BH466685 (1-797)
 Qy 13 ValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAspThrGlnLeuGluGlu 32
 Db 90 GTCACCAAGGCACCTAGACGGCGCGTTCACCTTTCCGAGCCAGCTACTCAAGATCGGTAG 149
 Qy 33 SerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeu 52
 Db 150 GCCATGAGGTATAGCTTCTCTCAACGCGCAACGCGTAAGCAATGCTTTGTTGGCT 209
 Qy 53 ThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGlu 72
 Db 210 GCATGCGAGCTCGTGGCGGCGCATGATCAACCCCAATGCTTTGCTGTGCTTCCAG 269
 Qy 73 MetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAspAspTyr 92
 Db 270 ATGATCCACGCGTCTGCTATCATCTTAGACCACTCCCTTCATGGAGGAGCAGCAGC 329
 Qy 93 ArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluThrPThrAlaIleLeuAlaGly 112
 Db 330 CGCGCGGAAAAACCCACCACCACTAATATTTGCGAGGATAGCGCGGTTTGGCTCT 389
 Qy 113 AspAlaLeuLeuThrLysAlaPheGluLeuLeuSerSerAspArgLeuThrAspGlu 132
 Db 390 GTTCGCTCATAGTTGGCGCTCAACCAACCTCTTCATCCCGGATGTTCTTCCACAG 449
 Qy 133 ValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGly 152
 Db 450 AGATTCTCCGGCGGTTTCAGTGTATGATGCGAGCGGTGGGACAGAGAGGACTCGTCGG 509
 Qy 153 GlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp-----Leu 167
 Db 510 GGACAAGCGGAGGATCTTGGCGGAGAGGATGAGCTTCGAGGAAAAATGAAGCGCGCTG 569
 Qy 168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
 Db 570 GAACATCTTGATTTATACATATTCATAAAACGCGCTCGCTTGAAGCTGCTGCGGTG 629
 Qy 188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
 Db 630 GTGGGACCAATAATGGGAGGTGGCTCTCATGAGGAGATGAAGAGCTTGGGAAATACGCG 689


```

Db 298 CGGCCATTGGTGAGCTCGCGCGCTGTCATCGGATCAATAGGCGCTCGTTGGCGGTACGGTT 357
Qy 156 LeuAspMetGlnSer-----GluGlyGlnProIleAspLeuThrLeuGluMetIle 173
Db 358 GTTGATCTGGAGTACTGTCTCACTGACCTGTACCTTACCGCCCTGGAGTACATT 417
Qy 174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAla 193
Db 418 CATGTGCACAAGACCGCTGCTGCTGAGGCTCTCGGGTTATTGGTGCAATCATTCGG 477
Qy 194 AsnValAspAspThrLysGluHisLeuGluSer-----TyrSerTyrHisLeu 210
Db 478 GGTGGCTCAGAC-----GAGCACATCGAGCAGGTGCTCATGTACGCCAGATCACTC 528
Qy 211 GlyMetMetPheGlnIleLysAspLeuLeuAspCysTyrGlyAspGluAlaLysLeu 230
Db 529 GAGCTGCTCATCGCGCGATGACGATCCTTGATGTACCAAGTATCAGATGAGCTA 588
Qy 231 GlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeu 249
Db 589 GGGAGACAGCTGTGTAACGACTTGGCGAGTGACAGACCCCATACCCCATTTACTG 645

```

RESULT 10

BH250143/c
LOCUS BOGAG35TF BOGA Brassica oleracea genomic clone BOGAG35, DNA 770 bp 26-NOV-2001

DEFINITION BH250143

ACCESSION BH250143

VERSION BH250143.1 GI:17073585

KEYWORDS GSS.

SOURCE Brassica oleracea.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 770).

Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished (2001)

Other GSSs: BOGAG35TR

Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

FEATURES

source

1..770
/organism="Brassica oleracea"
/strain="T01000PH3"
/db_xref="taxon:3712"
/clone="BOGAG35"
/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 8,47E-33 Length: 770
Score: 339.50 Matches: 80
Percent Similarity: 62.23% Conservative: 37
Best Local Similarity: 42.55% Mismatches: 58
Query Match: 23.37% Indels: 13
DB: 17 Gaps: 5

US-09-925-637-64 (1-287) x BH250143 (1-770)

Qy 6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25

```

Db 539 GTGAACAAGACTCTAGACGAA-----GCCATACAGTTCGGGAGCCACTC 495
Qy 26 MetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
Db 494 -----AAGATCCAGAGCCATCGTTACGCGTTCTTCGACCGCGGAAACGGTGA 444
Qy 46 ArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
Db 443 AGGCCATTCTTCGCTCGCTTCTTGGAGCTAGTAGCGCGGCGAAGAACCGCGCATG 384
Qy 66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
Db 383 CCAGCGGCTTGTGCGGTTGAGATCATACACACCATGTCTCTTATCAAAAGCAGCTTCC 324
Qy 86 AlaMetAspAsnAspAspTyrArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
Db 323 TGCATGACATGACGACCTGCGTGGGAGGCCACCACTCACAAAGCCTTCGGCGAA 264
Qy 106 TrpThrAlaIleLeuAlaGlyAspAlaLeuThrLysAlaPheGluLeuIleSerSer 125
Db 263 GGAATCGCCATTCTCCCGGAGGAGCTCTCTTGTCTCTTGGTTCGACATGACACAG 204
Qy 126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
Db 203 GCTGAG---ATATCTCGGAGAGAAATGGTTGGCGGCTCAGGAACCTGGCTAGTCCAT 147
Qy 146 GlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGly----- 162
Db 146 GAACTAGAGGGTGGTGGCGGAGCAAGCCAAAGACATAGTAGTGAAGGTTTGGAGTTG 87
Qy 163 GlnProIleAspLeuThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeu 182
Db 86 AACGAGCTCGGTTAGGACATTTGGAGTTATCCATCTACACAAACCGCTGTTTGTG 27
Qy 183 ---ThrPheAlaValMetSerAla 189
Db 26 GAAACTGCTGCGGTTCTTGGAGCC 3

```

RESULT 11

AJ302129

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

PLANT BIOLOGY

FRIEDRICH MIESCHER-INSTITUT

MAULBEERSTRASSE 66, 4058 BASEL, SWITZERLAND.

LOCATION/QUALIFIERS

1..442

source

/organism="Solanum tuberosum"

/cultivar="Bintje"

/db_xref="taxon:4113"

/clone="S3-2B"

/tissue_type="leaf"

/note="infected with Phytophthora infestans"

BASE COUNT

ORIGIN

104 a 102 c 125 g 111 t

AJ302129 442 bp mRNA linear EST 03-APR-2001
clone S3-2B, mRNA sequence.
Solanum tuberosum cv. Bintje leaf

GI:13536952

potato.

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 442)

Beyer, K., Binder, A., Bollner, T. and Collinge, M.A.

Identification of potato genes induced during colonization by

Phytophthora infestans

Unpublished (2001)

Contact: Beyer K

Plant Biology

Friedrich Miescher-Institut

Maulbeerstrasse 66, 4058 Basel, Switzerland.

Location/Qualifiers

1..442

/organism="Solanum tuberosum"

/cultivar="Bintje"

/db_xref="taxon:4113"

/clone="S3-2B"

/tissue_type="leaf"

/note="infected with Phytophthora infestans"

BASE COUNT

ORIGIN

104 a 102 c 125 g 111 t

Score: 299.00 Matches: 64
Percent Similarity: 68.91% Conservative: 18
Best Local Similarity: 53.78% Mismatches: 29
Query Match: 20.58% Indels: 8
DB: 13 Gaps: 2
US-09-925-637-64 (1-287) x BI717621 (1-666)
QY 6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
DB 331 GTGAACAAGGCCATGACGAG-----GCCGTGCCGTGAAGTAC 369
QY 26 MetAspThrGlnLeuGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
DB 370 CCGGAGACT---CTGAACGAGACATGCGCTACTCCTGCTGGTGGCGCAAGCGCTG 426
QY 46 ArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
DB 427 CGCCAGCTCTCTGCTGGCGGCTGGAGCTGGTGGCGGACATCCACGCCGCGCTG 486
QY 66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspLeuPro 85
DB 487 CCTGCTGGTGGCGCATGGAGATGGTGCACACCATGCTCATCATGACGACCTGCC 546
QY 86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
DB 547 TCCATGGACAACGACGACTTCGCGCGCGCGCCGCCAACCAACCAAGGTGTACGCGGAG 606
QY 106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer 124
DB 607 GACATTGCCATCTCTGGCGGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663

Search completed: May 30, 2003, 21:50:33
Job time: 1562 secs

310 CCGCGCGCTCGCGGCGAGATGATTCACACGATGCTACATCAAGACGATCTCCT 251
QY 86 AlaMetAspAsnAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
DB 250 TGCATGACACGACGACCTCCGCGAGGTAAGCCAGGACACAAAGTCTTCGGAGAA 191
QY 106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer 125
DB 190 AGCGTCGCGATCTCTCCGCGCGCGCTCTAGCTCTGGCTTCGAGCATTTGACGGAA 131
QY 126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
DB 130 GCTGAC---GGTCCGCTGAGAGATGTTAGCGGTTAAGAACTGCGGAAGTCTATA 74
QY 146 GlyHisValGlyMetValGlyGlnMetLeuAspMetGlnSerGluGly 162
DB 73 GGGAGAAAGGCTCGTGGCGGACAAAGCGATGTTGACGAGGAGGT 23

RESULT 15
BI717621
CUS
DEFINITION 666 bp mRNA linear EST 19-SEP-2001
Lambda Zap II C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
BI717621
BI717621.1 GI:15693316
EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
vascular Plants. Project: 1031
JOURNAL Unpublished (2001)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
FEATURES
source
1. .666
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
) , Lambda Zap II"
/note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."

BASE COUNT 126 a 211 c 197 g 132 t
ORIGIN
Alignment Scores: 1.05e-27 Length: 666
Pred. No.:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 17:35:15 ; Search time 2483 Seconds
(without alignments) : 10091.629 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atacgaatcaccgatgaa.....tattagaatcgttgattta 861

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

d size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:**

- 1: gb_ba:**
- 2: gb_htg:**
- 3: gb_in:**
- 4: gb_om:**
- 5: gb_ov:**
- 6: gb_pat:**
- 7: gb_ph:**
- 8: gb_pl:**
- 9: gb_pr:**
- 10: gb_ro:**
- 11: gb_sts:**
- 12: gb_sy:**
- 13: gb_un:**
- 14: gb_vi:**
- 15: em_ba:**
- 16: em_fun:**
- 17: em_hum:**
- 18: em_in:**
- 19: em_mu:**
- 20: em_om:**
- 21: em_or:**
- 22: em_ov:**
- 23: em_pat:**
- 24: em_ph:**
- 25: em_pl:**
- 26: em_ro:**
- 27: em_sts:**
- 28: em_un:**
- 29: em_vi:**
- 30: em_htg_hum:**
- 31: em_htg_inv:**
- 32: em_htg_other:**
- 33: em_htg_mus:**
- 34: em_htg_pln:**
- 35: em_htg_rod:**
- 36: em_htg_man:**
- 37: em_htg_vrt:**
- 38: em_sy:**
- 39: em_htgo_hum:**
- 40: em_htgo_mus:**
- 41: em_htgo_other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	810	94.1	333750	1	AP004827 Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
C 2	657	76.3	301550	1	AP003134 Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
C 3	657	76.3	346900	1	AP003362 Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
C 4	606	70.4	882	6	AR106466 Sequence
C 5	25	2.9	783	6	AX141501 Sequence
C 6	25	2.9	3444	1	AF270104 Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
C 7	25	2.9	3444	6	AX145422 Sequence
C 8	25	2.9	3454	1	AF269958 Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
C 9	25	2.9	3454	6	AX145276 Sequence
C 10	25	2.9	4045	1	AF269889 Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
C 11	25	2.9	4045	6	AX145207 Sequence
C 12	25	2.9	111507	9	HS238017 Sequence
C 13	22	2.6	227165	2	AC096106 Sequence
C 14	21	2.4	135599	8	CPU30821 Sequence
C 15	21	2.4	157270	9	AF236874 Homo sapiens
C 16	21	2.4	158344	10	AL133401 Mouse DNA
C 17	21	2.4	159958	9	AF000475 Homo sapiens
C 18	21	2.4	170630	2	AC034163 Homo sapiens
C 19	21	2.4	171252	2	AC129764 Rattus norvegicus
C 20	21	2.4	172663	9	AF241725 Homo sapiens
C 21	21	2.4	180650	9	AC109474 Homo sapiens
C 22	21	2.4	200029	2	AC101844 Mus musculus
C 23	21	2.4	340000	9	HS21C027 Homo sapiens
C 24	20	2.3	508	10	AF352782 Mus musculus
C 25	20	2.3	884	10	MUSSPKC
C 26	20	2.3	3008	5	AF254639 Danio rerio
C 27	20	2.3	3282	1	STVPCDEF
C 28	20	2.3	4942	10	S48768 SP-A-pulmon
C 29	20	2.3	10029	1	AE006320 Lactococcus lactis
C 30	20	2.3	22779	1	AE008791 Salmonella enteritidis
C 31	20	2.3	29303	3	CEF57G4
C 32	20	2.3	32612	3	CET05D4
C 33	20	2.3	70446	9	AC107299 Homo sapiens
C 34	20	2.3	103308	9	AC092491 Homo sapiens
C 35	20	2.3	105208	10	AL671897 Mouse DNA
C 36	20	2.3	118135	9	AC117477 Homo sapiens
C 37	20	2.3	126117	2	AL157949 Homo sapiens
C 38	20	2.3	126954	9	AL365440 Human DNA
C 39	20	2.3	129584	2	AC101911 Mus musculus
C 40	20	2.3	140765	9	AC021301 Homo sapiens
C 41	20	2.3	146058	9	AC003668 Homo sapiens
C 42	20	2.3	149777	2	AP003573 Oryza sativa
C 43	20	2.3	153369	9	AL592047 Human DNA
C 44	20	2.3	156305	2	AC120452 Rattus norvegicus
C 45	20	2.3	158151	9	CNS05TDT Human chromosome

ALIGNMENTS

RESULT 1
AP004827/c
LOCUS
DEFINITION
Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA, complete genome,
strain:MW2, section 6/10.
ACCESSION
AP004827 BA000033
VERSION
AP004827.1 GI:21204509
KEYWORDS
SOURCE
Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
ORGANISM
Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus
REFERENCE
AUTHORS
1
Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
Yamamoto,K. and Hiramatsu,K.

AP004827 333750 bp DNA linear BCT 02-JUL-2002
Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
strain:MW2, section 6/10.
AP004827 BA000033
AP004827.1 GI:21204509

TITLE Genome and virulence determinants of high virulence community-acquired MRSA
JOURNAL Lancet 359 (9320), 1819-1827 (2002)
MEDLINE 22040717
PUBMED 12044378
REFERENCE 2 (bases 1 to 333750)
AUTHORS Aoki, K., Oguchi, A., Nagai, Y., Asano, K., Iwama, N., Baba, T., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) Akio Oguchi, National Institute of Technology and Evaluation, Biotechnology Center; 2chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail: oguchi@nitech.go.jp, URL: http://www.bio.nite.go.jp/, Tel: 81-3-3481-8423, Fax: 81-3-3481-8424)
FEATURES
Source Location/Qualifiers
1. 333750
/organism="Staphylococcus aureus subsp. aureus MW2"
/strain="MW2"
/db_xref="taxon:196620"
complement(278..619)
/gene="MW1341"
complement(278..619)
/gene="MW1341"
/note="ORFID: MW1341"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB95206.1"
/db_xref="GI:21204510"
/translation="MTNITKVFQDNLEQIDVFANKVLGSSSEAKPYLDEYHSKIIDI
FYQDRLTGNDIFDKLQNPVPMNFKERYMIERYHFMGYRQMKTKTELKMA
SQTQLNKQV"
complement(624..1283)
/gene="nth"
complement(624..1283)
/gene="nth"
/codon_start=1
/transl_table=11
/product="endonuclease-like protein"
/protein_id="BAB95207.1"
/db_xref="GI:21204511"
/translation="MYSKKALEMIDIVIANPMPDAECLEKHNPELTIAVLLSAQCT
DYLNRVTELFKTKTPTYDLAVSDEELMNDIRSIGLYRNKAKNKKLCQSLDQFN
GIPOTHELESAGVQKTNVMSVAFDEPLAVDTHVERVSRGLGNRWKDNVRQ
VBDRLCSVIPDRWNKSHQLFFGCRYHCLARKPKDCICPLEDCREGQRYKYSKE
A"
complement(1273..1959)
/gene="MW1343"
complement(1273..1959)
/gene="MW1343"
/note="ORFID: MW1343"
hypothetical protein, similar to chromosome replication
initiation protein dnad"
/codon_start=1
/transl_table=11
/protein_id="BAB95208.1"
/db_xref="GI:21204512"
/translation="MDKYQKAPVPIRELLDHYSDLDQDLVILLKLIYASETS
NKPSIELQKSTQWQPDITNVIONLQRELLQVQNDREGETEYMLDPFFKEL
SHLQSMETKEQSKKFKOLFVLEDTFARPLSPYEITFLNQMIDVDKHSIIQ
WALDANSILNKLSFYMDRILLNWKNNVKTIDDSRKIREKFNKPKMTHTKVTPKFD
WNGSNLDGK"
complement(2287..3579)
/gene="asnS"
complement(2287..3579)
/gene="asnS"
/note="ORFID: MW1344"
/codon_start=1
/transl_table=11
/product="asparaginyl-tRNA synthetase"
/protein_id="BAB95209.1"
/db_xref="GI:21204513"
/translation="MKTTIKQAKDHLNQDVTIGAWLTNKRSSGKIAFLQLRDGTGMQ
GVVKSVEDEEYKLAKEITQESSLYVTGTTEDNRSGLGYEMQVKSIEVISEADHYD
TTPKNGTEFLMDHRLHLSRKQHAVMKIRNEVIRATYEFENQDGTGVTPIILTAS
APECTSELPHTKYFDODAFLSQSGOLYLEAAMAKGVFSFGTTPRAESKSRHLLIE
FMWIGEMAFTHAESLEIQEQTIVHVVSLECKLEKILERTDITKIKKVPMPFR
ISYDDAEFLKSEGDDEWDFGAFETAIANHYDLEVFITNTPTKIKPYPMPNP
ENEETVGLADIAPEGYGEIIGSERVDDLELEQVRKEHGLDEEAYSYYLDLRRYS
VPHCGGLCLERTVAWISGVEHVRETAPPPLRLNRLYP"
complement(3901..6594)
/gene="ding"
complement(3901..6594)
/gene="ding"
/note="ORFID: MW1345"
/codon_start=1
/transl_table=11
/product="probable ATP-dependent DNA helicase ding"
/protein_id="BAB95210.1"
/db_xref="GI:21204514"
/translation="MGMATYAAVVDLETTGNQLDFFDIIQIGITFVRNNOIIDYHSMI
RTNLETPPFIQALTSIEENMLQOAPYFNQVQAEIYDKIKDCIFVAHNVDFDLNFKKA
PKDCNIOVRPKVIDTLEIFKTAFTDKSYQLSELAEAHGITLANAHRADDAATTA
LMLAFEFKELPLDTLQLYSLKOLYLDYIFEEVVRQYDAKPLDSKYEFGQII
YKQVDFKPTTNYNGSLSKSAVDQGLTYRQQLYLAETILDQLMHSKAMTEA
SIGSGSLAYLLAALMINIETKHMVNTSTKLQSLQLEKIDIPAMNEALNFKINALL
IKSDYSISGLISQILKDDTSNYEVNLMQLLITWITETPSGDIQELNLKGQKMYF
DQKIETVYPARDHVHYNFIRKNAQIQIGITNHAHLIHSVDENSYQLFDDCIVDEA
HRLPYALNOVTNELSYADIKYQGLIGKNENEKLLKADQLEKORILEKLDIAPIDI
FGLKASMEIHELNEQLESTIFTIINDSDVYDDDIHRFHNVTFFETKDLKDLHAID
KLKLEIYFNGISHTKTVKSLRQLYLKDKFNIEQSLKAGHTSISIKNLQSKSTIR
LVYDQAVKDLQTVLEKFLSLIFISGLTKFNHSEFAFKFNKDVHTEFEVNTSL
QSAKTSVIFPSDVASYQKNIDEVASIVYIEYTTTSKSLVLTFTSYKMHMVQ
DMNELPEFEDVYLTQQONKYKTVQFNEDKAILLGTSTFFEGDFQANGIKVM
IAKLPMKNHAKYWLMDSEFTSTFEYVLPDAVTRFQGLRGLRSENDRGIIIVSFD
DRLINSYKNFEPTLENYROAKGDIQPGKLLRQIQKKK"
complement(6618..7589)
/gene="MW1346"
complement(6618..7589)
/gene="MW1346"
/note="ORFID: MW1346"
/codon_start=1
/transl_table=11
/product="truncated hypothetical protein"
/protein_id="BAB95211.1"
/db_xref="GI:21204515"
/translation="MSKYQDVLQLLYKNKPNYISQSIASISINTSRATVKKVIDQLK
LSECKIDSNHKGHLLOLPDIWYOGIDQYTKSALPDFSEVYDSIDSTQLAAKSL
VGNQSSFEPLTDEQTKGRFNRHSSSGQGLWMSVLRPNVAFSMISKFNLFALG
IRDAIQHSQDEVKVPNDIYIDNGKVCGLFTEWANNNDGEALICGIGNLTOOLE
NFEISIRHRATSVQLHDKNKLDRYQFLERLQGEIEKRYNQFLTLPFSEIREYIAASN
IWNRTLFTENDQKFGAIDLDYDGYLIVRDEAGESHRLISADIDF"
complement(7576..8778)
/gene="MW1347"
complement(7576..8778)
/gene="MW1347"
/note="ORFID: MW1347"
hypothetical protein, similar to poly(A) polymerase"
/codon_start=1
/transl_table=11
/protein_id="BAB95212.1"
/db_xref="GI:21204516"
/translation="MDKSLFEQARPILQIQDNGFEAYVVGVSVDYVGNRNIHIDI
TTSAPDEIESIFSHITIPVKEHGNTINVYNDENVTFTFRAEEDYVDRHPSGVTFF
RDLYEDLQRDRFTMAIAMDYKLYDFDQDIDNNRIIRTVGTAERFQDGLRMI
RCLRFQSLSDFIATETPEAMRIQADIKFLSILRIYELTKLMRGINVERSFNHLKS
LKAFNMPYFEHLDNQINVTAEIOLLELIAIVSVKFDINYSLEKPLKLSNRQVKDINQ
YIQIMNALPSIITKEQLKMFYDIDTHLLIKNVVAAVDYKANDIQGHEPLIVNLQDID
ETLHRLPMHNRKMDVMVNGVLMHLNAKSGPWLKQVLRQIEIAIVTGVKVSNEETELK
WYDNIHVKI"
complement(8783..9925)
/gene="MW1348"
complement(8783..9925)
/gene="MW1348"
/note="ORFID: MW1348"

hypothetical protein, similar to lipopolysaccharide biosynthesis-related pr homolog
/codon_start=1
/transl_table=11
/protein_id="BAB95213.1"
/db_xref="GI:21204517"
/translation="MKGITCPSMGSGGIATTELKLAERGHEVHITSNIPRIR
APLPMFIHQVEVNOYAVQPPYDITLSTKIAEVIKEYGLDLHMHYAVPHAICGIL
AREMSGKIDKIMTTLHGDTITVLGSDHSAIKFGEKSDIVTSVSKLAQETHIIEI
ETNEKIDPIYFNVENFPTKHTALAKSQFIAPDEKVLHIVSNFROVKRIDTIIET
AKREKIPSKILLGDPGLPMRLKELNVEDVFLGKODCVSEYOLSDVLVLL
SEKSEFGLTLEAKMTGVPIGNSNAGIKVKKIYKGTGTFVVDGDCSASDYAIRLLE
DKVLYNKLQKWLADIARFEGSELITDQYEEYQKMLNEHNSKGE"
complement(10167..10484)
/gene="MW1349"
CDS
complement(10167..10484)
/gene="MW1349"
/note="ORFID:MW1349"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95214.1"
/db_xref="GI:21204518"
/translation="MKSVMEMOREDEVIGOFKGYEPLANLRLTEEVGELAREIN
HTYGEKKKDEEANTKAEGLDNLFLVLLCLANSMDMTESFNETMEKFNTRKNEF
ERK"
complement(10820..11518)
/gene="MW1350"
CDS
complement(10820..11518)
/gene="MW1350"
/note="ORFID:MW1350"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95215.1"
/db_xref="GI:21204519"
/translation="MSSIFTMIIYFVILVPLNQHVKVSEYKSVYTSKTKGR
EVAIEHLANGIYDEVYKGEGLFDHPKRVYVLSVPANDRPSVAGTALIAHVEG
HAIQDHQWFLRFVAAVLPVANGSSLSYIMILGIIITLTAIGSFASTALWICAGLM
SLAVLFSITVLPVDFDASSRAKQITALNIVNEREYKHARKVLSAAAAMYVAATAVAV
AELVRIILLARSSD"
Query Match 94.1%; Score 810; DB 1; Length 333750;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 860; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGACGAATCTACCGATCAATAAATTAATAGATGAAGTCAATAATGAATTCGGTTGCG 60
118374 ATGACGAATCTACCGATCAATAAATTAATAGATGAAGTCAATAATGAATTCGGTTGCG 118315
QY 61 ATAAATAATCAGTAATGGATCTAGCTAGAGAAAGATATGTTGTTATTCATTAATGCT 120
118314 ATAAATAATCAGTAATGGATCTAGCTAGAGAAAGATATGTTGTTATTCATTAATGCT 118255
QY 121 GGAGTAAACCCATCCGACCACTGCTGTATTACTCATTCTAGATTCTACTAAATACCCAG 180
118254 GGAGTAAACCCATCCGACCACTGCTGTATTACTCATTCTAGATTCTACTAAATACCCAG 118195
QY 181 TATGAGTTAGGTATGAAGAGCGCAATTCACATAGAAATGATTATACATATTCATTATT 240
118194 TATGAGTTAGGTATGAAGAGCGCAATTCACATAGAAATGATTATACATATTCATTATT 118135
QY 241 CATGATGACCTACCGAGCGATGATTAATGATGATTATCAGCAGGAGAAATTAACAATCAT 300
118134 CATGATGACCTACCGAGCGATGATTAATGATGATTATCAGCAGGAGAAATTAACAATCAT 118075
QY 301 AAAGTATATGTTGAGTGACCTGCGATATACAGGTGATGCTTTATTAACHTAAAGCATTT 360
118074 AAAGTATATGTTGAGTGACCTGCGATATACAGGTGATGCTTTATTAACHTAAAGCATTT 118015
QY 361 GAACATTTTCAAGTGATGATGATTAATCATGAAGTAAATAAATAAAGTTCTACAACGG 420
118014 GAACATTTTCAAGTGATGATGATTAATCATGAAGTAAATAAATAAAGTTCTACAACGG 117955

QY 421 CTGTCAATAGCAAGTGGTCATGTTGGAATGTCGGCGGTCAAATGTTAGATATGCAAAAGC 480
117954 CTGTCAATAGCAAGTGGTCATGTTGGAATGTCGGCGGTCAAATGTTAGATATGCAAAAGC 117895
QY 481 GAAGCCCAACCAATTTGATCTTGAACCTTTGGAATGATACACAAAACAAAACAGGAGCA 540
117894 GAAGCCCAACCAATTTGATCTTGAACCTTTGGAATGATACACAAAACAAAACAGGAGCA 117835
QY 541 TTATTAACCTTTTGGCGTTATGATGTCAGCAGATATCGCTAAATGTCGATGATCAACTAAA 600
117834 TTATTAACCTTTTGGCGTTATGATGTCAGCAGATATCGCTAAATGTCGATGATCAACTAAA 117775
QY 601 GAACATTTAGAAAGTTATGTTTATCATTTAGTATGATGTTTCCAGATTAAAGATGATTTA 660
117774 GAACATTTAGAAAGTTATGTTTATCATTTAGTATGATGTTTCCAGATTAAAGATGATTTA 117715
QY 661 TTAGACTGCTATGTTGATGAAGCAAGTTAGTAAAGAGTGGGAGCGATCTTGAANAAT 720
117714 TTAGACTGCTATGTTGATGAAGCAAGTTAGTAAAGAGTGGGAGCGATCTTGAANAAT 117655
QY 721 AATAAAGTACGTACGTGAGTTTATAGGAAAGATGGCGCAAGATGATGATGATGATGAT 780
117654 AATAAAGTACGTACGTGAGTTTATAGGAAAGATGGCGCAAGATGATGATGATGATGAT 117595
QY 781 CATAGAGCGCAGCAGTGGATGAACATAACGCAAAATTTGATGAACAAATTCATCAACAAACAC 840
117594 CATAGAGCGCAGCAGTGGATGAACATAACGCAAAATTTGATGAACAAATTCATCAACAAACAC 117535
QY 841 TTATTAGAAATCGTTGATTTA 861
117534 TTATTAGAAATCGTTGATTTA 117514
Db
RESULT 2
AP003134/C
LOCUS AP003134 301550 bp DNA linear BCT 02-JUL-2002
DEFINITION Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
genome, section 6/10.
ACCESSION AP003134 BA000018
VERSION AP003134.2 GI:14349226
KEYWORDS
SOURCE Staphylococcus aureus subsp. aureus N315 (sub_species:aureus N315,
strain:N315) DNA.
ORGANISM Staphylococcus aureus subsp. aureus N315
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
Sekimizu,K., Hirakawa,H., Kihara,S., Goto,S., Yabuzaki,J.,
Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
TITLE Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
JOURNAL Lancet 357 (9264), 1225-1240 (2001)
MEDLINE 21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 301550)
AUTHORS Aoki,K., Oguchi,A., Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K.
and Kikuchi,H.
TITLE Direct Submision
JOURNAL Submitted (30-JAN-2001) Akio Oguchi, National Institute of
Technology and Evaluation, Biotechnology Center, 2Chome 49-10
Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/
Tel:81-3-3481-8423, Fax:81-3-3481-8424)
COMMENT On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES
Location/Qualifiers
source
1..301550
/organism="Staphylococcus aureus subsp. aureus N315"
/strain="N315"


```
/protein_id="BAB42562.1"
/db_xref="gi:13701268"
/translation="MERTFLIKPDVAVORNLIIGVISRIERKGLKLYGGKLMQVPMEL
ATHYGHQGFNFYDLISFTISAPVFAMVVEGDVAVNSRHIIIGSNPSSESPGSR
ADGLTVGRNIIHGSDSLESAREINLWFNENEITSVAPRDWLYE"
complement(8318..9277)
/gene="gerCC"
complement(8318..9277)
/gene="gerCC"
/feature="ORFID:SA1302"
/codon_start=1
/transl_table=11
/product="heptaprenyl diphosphate synthase component II"
/protein_id="BAB42563.1"
/db_xref="gi:13701269"
/translation="MAKLNMMNEIKKVEORLEKAKSKDSVLEQASHLHSSGKRRV
PAFVILSQFGKDEQTSQTYQVAVALLEIHMATLVHDDVIDKSKRRGKLTISKWD
```

Query Match 76.3%; Score 657; DB 1; Length 301550;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 857; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	1	ATGACGAATCACCAGTGAATAAATTAATAGATGAAGTCAATAATGAATATCGGTTCCG	60
DB	64130	ATGACGAATCACCAGTGAATAAATTAATAGATGAAGTCAATAATGAATATCGGTTCCG	64071
QY	61	ATAAATAATCAGTAATGGATCTACAGTAGAAGAAAGTATGTTGTTATCTAATAATGCT	120
DB	64070	ATAAATAATCAGTAATGGATCTACAGTAGAAGAAAGTATGTTGTTATCTAATAATGCT	64011
QY	121	GGAGTAAACCCATCCGACCACTCTGTTATCTACTCTTATGATCTAATAATCCGAG	180
DB	64010	GGAGTAAACCCATCCGACCACTCTGTTATCTACTCTTATGATCTAATAATCCGAG	63951
QY	181	TATGAGTTAGTATGATGAGCGCAATTCACATAGAAATGATTCATATTTCACTTATT	240
DB	63950	TATGAGTTAGTATGATGAGCGCAATTCACATAGAAATGATTCATATTTCACTTATT	63891
QY	241	CATGATGACCTACCAGCGATGGTAATGATGATTATGACGAGGAAATTAACAATCAT	300
DB	63890	CATGATGACCTACCAGCGATGGTAATGATGATTATGACGAGGAAATTAACAATCAT	63831
QY	301	AAAGTATGATGATGAGTGCATTCATGATGATGATGATGATGATGATGATGATGAT	360
DB	63830	AAAGTATGATGATGAGTGCATTCATGATGATGATGATGATGATGATGATGATGAT	63771
QY	361	GAACCTATTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	420
DB	63770	GAACCTATTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	63711
QY	421	CTGTCAATAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	480
DB	63710	CTGTCAATAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	63651
QY	481	GAAGCCCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	540
DB	63650	GAAGCCCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	63591
QY	541	TTATTAACCTTTTGGCGTTATGATGATGATGATGATGATGATGATGATGATGATGAT	600
DB	63590	TTATTAACCTTTTGGCGTTATGATGATGATGATGATGATGATGATGATGATGATGAT	63531
QY	601	GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660
DB	63530	GAACATTTAGAAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	63471
QY	661	TTAGACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	720
DB	63470	TTAGACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	63411
QY	721	AATAAAGTACGCTACGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
DB	63410	AATAAAGTACGCTACGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	63351

QY	781	CATAGAGCGCAGCAGTGGATGAACGCAATTAATGATGAACAAATTCATTAACAAACAC	840
DB	63350	CATAGAGCGCAGCAGTGGATGAACGCAATTAATGATGAACAAATTCATTAACAAACAC	63291
QY	841	TTATTAGAAATCGTTGATTTA	861
DB	63290	TTATTAGAAATCGTTGATTTA	63270
RESULT 3			
AP003362/c			
LOCUS	346900 bp	DNA	linear BCT 07-FEB-2002
DEFINITION	Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete		
sequence, section 5/9.			
AP003362 BA000017			
ACCESSION	AP003362.2	GI:14247083	
VERSION			
KEYWORDS			
SOURCE	Staphylococcus aureus subsp. aureus Mu50 (sub_species:aureus Mu50, strain:Mu50) DNA.		
ORGANISM	Staphylococcus aureus subsp. aureus Mu50		
REFERENCE	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
AUTHORS	Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hoshino, A., Mizutani-Ui, Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.		
TITLE	Whole genome sequencing of methicillin-resistant Staphylococcus aureus		
JOURNAL	Lancet 357 (9264), 1225-1240 (2001)		
MEDLINE	21311952		
REFERENCE	2 (bases 1 to 346900)		
AUTHORS	Ohta, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College of Medical Technology and Nursing, Department of Medical Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan (E-mail:tohta@akura.cc.tsukuba.ac.jp, Tel:81-298-53-3454, Fax:81-298-53-3454)		
COMMENT	On May 29, 2001 this sequence version replaced gi:13875626.		
FEATURES	Location/Qualifiers		
source	1. 346900		
	/organism="Staphylococcus aureus subsp. aureus Mu50"		
	/strain="Mu50"		
	/sub_species="aureus Mu50"		
	/db_xref="taxon:158878"		
gene	12..263		
CDS	12..263		
	/gene="SAV1313"		
	/gene="SAV1313"		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="BAB57475.1"		
	/db_xref="GI:14247084"		
	/translation="MLTLIKLDEQVLIIEYIPEDDISNGKGSYTFNKKDAEVIDFS		
	LSSENEYEFMLYRNKSFVSVRDFIEKQEPENKIANY"		
gene	987..1181		
	/gene="SAV1314"		
CDS	987..1181		
	/gene="SAV1314"		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="BAB57476.1"		
	/db_xref="GI:14247085"		
	/translation="MNLQSYIGKLVKLTLTNKILIGKLVDFDDRVDFDGVNSEID		
	TGRITYDISENKKITIVLOK"		
gene	complement(1725..2750)		
	/gene="SAV1315"		

```
CDS
complement(1725..2750)
/gene="SAV1315"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB57477.1"
/db_xref="GI:14247086"
/translation="MISFENDYLEGAHEKVLKRLVTLNVLQASGYGDFQTAQAIEKI
KDTIDGPNATIRLVGGTQNTQVIVINSMLSEYEGVISADTGHVAVHEGGAIEYSGHKV
ITIPSEKVASDSVETMETFKSDFKDHVPFGMVYISHPTVEYGTLYKSELELC
KVKQYQVLPFMDGARLGVLGMSQSDMTIKDKAKYCDVFEYIGGTIGALCGEAIYFT
KNNEKQFTRIKHGALLAKLGTGIELEFDNLNINISRHATEMANKMGDKGFKN
KGRYLFDSPTNQFFILSNEKTAELEQKVFVWEKVIDDQHRVVPATSWATTEENL
NKLELEI"
3020..3217
/gene="SAV1316"
3020..3217
/gene="SAV1316"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57478.1"
/db_xref="GI:14247087"
/translation="MQQIKLFTFTAELELLEKLNINAPLSSEATNKLKLVNITIKIEI
ERTFPNNEEFNAILLSVKN"
3238..4755
/gene="SAV1317"
3238..4755
/gene="SAV1317"
/codon_start=1
/transl_table=11
/product="cardiolipin synthetase homolog"
/protein_id="BAB57479.1"
/db_xref="GI:14247088"
/translation="MEFVPEFFIGETMYTESNDLGTFTIILAIGFIINLVLAFTII
FLERNRTASSTWAWLFVLVPLIGLFLYLPFGVYSARKLNKNGNVLTFDGLLK
QIESPDKNYGTNDKQVQKHDLVSNLLMDQGFTEKNKVDHFIDGNDLYDQVLKD
INAKRIHLEYTFALDGLGRKHLALEKRLQGLEVKLYDDVSGKNVKNWAFDHF
KSLGVEAFASKDLNFRNNRHKIIVIDGOLYGVGGFNIGDEYLGKLGWY
RTHLIOGDADQLDRFLDWSQARPOEYDVKYFKKNGPLGNSPIQIAASGP
ASDWHQIEGYTKMISAKSVYLSQSPYIPDINSYINAKIAKSGVDVHLMPCKPD
HPLVWATPSNADSLSSGVKLYTVENGFIHSMCLIDDEIVSGVTANMDFSEFLNF
ENYAFVYDENLAKDLRAVEHDTIKSKQTKESYANRPLSVAFKESLAKLVSPIL"
4939..5436
/gene="SAV1318"
4939..5436
/gene="SAV1318"
/notes="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57480.1"
/db_xref="GI:14247089"
/translation="MQISNINKSPNKCVCVLANISFDIEQKCIALIGKAGKSTLI
DILIGNVANSEIFDKDLQSENRSINFQKTPPDQKLYEINLYQSFYENPLPL
EELIEITKDFSSQLNQFVNKLSGGQRLDLFVLSLIGQPQLILLDEPTSTMDIETREY
FWSII"
5461..5838
/gene="SAV1319"
5461..5838
/gene="SAV1319"
/notes="ATP-binding protein"
/codon_start=1
/transl_table=11
/product="ABC transporter homolog"
/protein_id="BAB57481.1"
/db_xref="GI:14247090"
/translation="MLTSHYLEEVRMSDKIILLENGEIILNDSTSHIRTNQSQSIT
LSDEYFKRLKLDKDLVLTQKNHNGTIKITSNVNDTILYLOQLHINLDIETQKRSIV
DSYFNKKQKQSGSNYDTKLEENRI"
5807..6538
/gene="SAV1320"
5807..6538

gene
/notes="similar to ABC transporter integral membrane
protein"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57482.1"
/db_xref="GI:14247091"
/translation="MILSYLKIEFKVIMRKTTLLLSILFPVIFYILFTSTLELPEDV
KPFYKEMYSMTVSYLSLSLTFEPLDIINEQNEQRQLMTWPTFTSYYSKVVYK
TMLQFAILYIVFWGHFYAGVAMSAVOWLESGIFLWLGASLLITFGILFSLNDIOK
TSALANIVTIGLWGLGFWPINTFPNMQHVARVLPYSYHLRKLGVDIAGNHHNINLIS
FATILLYVLSGSIIVYTCISFKRAE"
6542..7633
/gene="SAV1321"
6542..7633
/gene="SAV1321"
/notes="similar to two-component sensor histidine kinase"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57483.1"
/db_xref="GI:14247092"
/translation="MKPLKDTSAEISSILYLIPFIAGIFFNEVYGPKWLYIISVIVF
SLSYLILVNNRLNTLMFYILLIHYFIICYEYFVSHVPMLSLFFYSAPAVPTFKN
NVKKTATNFIITMTIITTYLYNNYFVAMVYVVISLIMLDFKMKMKREYQKE
IAEKRRHINTLIAEQERHRTIGDLHTDLGHVFASLSLKSLAYKLDTDVKYKAELL
AINKLSRESLNKVBRELIDVWKLPSFIEEDSIRKVLKADIDITFENKELAQVLSPK
QSMVITREAINNVINKHANSKRVHKLTVNNHKLMLIEDDDGKIDSDCEVKRSISQ
VOHLNGTLAVDSNTGTIIIEISTGGIA"
7630..8232
/gene="SAV1322"
7630..8232
/notes="similar to two-component response regulator"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57484.1"
/db_xref="GI:14247093"
/translation="MTSLIAEDQNMRLQAMVOLIKHGDFFELADTNDGLDAMKLE
EYNPNVVIDIEMPGMTGLEVAEIRKKHLNIRKIVITFRPGYFEKAVVNDVAYV
LERSLEELVETINKNGEYSATLMTSFFVDKPNLPKEQIVLREITNGLSLSEI
SEKLTLDGTVRNYSYIIDLKLFADNRFDAMKANKEGWI"
complement(8345..8533)
/gene="SAV1323"
complement(8345..8533)
/gene="SAV1323"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="MNSINVTIIVIIIGIIFLQIFFKLPLIVTAVLSILGI
FVGFIYLVISFYNNRKN"
8672..9205
/gene="nuc"
8672..9205
/gene="nuc"
/notes="SAV1324"
/codon_start=1
/transl_table=11
/product="thermonuclease"
/protein_id="BAB57486.1"
/db_xref="GI:14247095"
/translation="MKSNSLAMIVVAVIIIVGLAFQFMNHTGPFKKGTNHTVQDLN

Query Match 76.3%; Score 657; DB 1; Length 346900;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 857; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGACGAATCTACCGATGAATAAATTAATAGTAGTAATCAATGAATATCGGTGCG 60
|||||
```

Db 251159 ATGACGAATCTACCGATCAATAAATAATAGATGAAGTCAATAATGAATATATCGGTTGCG 251100
QY 61 ATAAATAATCAGTAATGGATCTACGCTAGAGAAAGTATGTTGTTATTCATTAATGCT 120
Db 251099 ATAAATAATCAGTAATGGATCTACGCTAGAGAAAGTATGTTGTTATTCATTAATGCT 251040
QY 121 GGAGTAAACGATCCGACCGAGTCTGTTATTTACTCAGTTTATGATTCATTAATACCGAG 180
Db 251039 GGAGTAAACGATCCGACCGAGTCTGTTATTTACTCAGTTTATGATTCATTAATACCGAG 250980
QY 181 TATGAGTTAGGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCATCTATT 240
Db 250979 TATGAGTTAGGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCATCTATT 250920
QY 241 CATGATGACCTACCGAGCGGATGATTAATGATGATTCATGACGAGGAAATTAACAAATCAT 300
Db 250919 CATGATGACCTACCGAGCGGATGATTAATGATGATTCATGACGAGGAAATTAACAAATCAT 250860
QY 301 AAAGTATATGTTGAGTGCAGTCCGATATAGCAGGTGATGCTTTATTAACATAAGCATTT 360
Db 250859 AAAGTATATGTTGAGTGCAGTCCGATATAGCAGGTGATGCTTTATTAACATAAGCATTT 250800
QY 361 GAACCTATTTCAGGTATGATGATTAATGATGATTAATGATGATTAATGATGATTAATGATGAT 420
Db 250799 GAACCTATTTCAGGTATGATGATTAATGATGATTAATGATGATTAATGATGATTAATGATGAT 250740
QY 421 CTGCAATPAGCAAGTGGTCTGTTGGAATGTTGCGCGGTCAAAATGTTAGATGATCAAAAGC 480
Db 250739 CTGCAATPAGCAAGTGGTCTGTTGGAATGTTGCGCGGTCAAAATGTTAGATGATCAAAAGC 250680
QY 481 GAAGCCCAACCAATTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTT 540
Db 250679 GAAGCCCAACCAATTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTT 250620
QY 541 TTATTAACCTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATCACTAAA 600
Db 250619 TTATTAACCTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATCACTAAA 250560
QY 601 GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAGATGATTTA 660
Db 250559 GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAGATGATTTA 250500
QY 661 TTAGACTCTATGTTGATGAAGCAAGTTAGTAAAGTGGCGAGCATCTTCAAAAT 720
Db 250499 TTAGACTCTATGTTGATGAAGCAAGTTAGTAAAGTGGCGAGCATCTTCAAAAT 250440
QY 721 AATAAAGTACGTACGTGAGTTTATTTAGGAAAGATGGCGAGAGATAAATTTGACTTAT 780
Db 250439 AATAAAGTACGTACGTGAGTTTATTTAGGAAAGATGGCGAGAGATAAATTTGACTTAT 250380
QY 781 CATAGAGCGCAGCAGTGGATGAACATAACGAAATTTGATGAACATAACATAACAAACAC 840
Db 250379 CATAGAGCGCAGCAGTGGATGAACATAACGAAATTTGATGAACATAACATAACAAACAC 250320
QY 841 TTATTAGAAATCGTTGATTTA 861
Db 250319 TTATTAGAAATCGTTGATTTA 250299

RESULT 4
AR106466
LOCUS AR106466 882 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6107058.
ACCESSION AR106466
VERSION AR106466.1 GI:12820996
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 882)
AUTHORS Gwyn.M. and Wilding,E.Imogen.
TITLE isPa from Staphylococcus aureus
JOURNAL Patent: US 6107058-A 1 22-AUG-2000;

FEATURES
source
BASE COUNT 328 a 118 c 186 g 250 t
ORIGIN
Query Match 70.4%; Score 606; DB 6; Length 882;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 856; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGACGAATCTACCGATCAATAAATAATAGATGAAGTCAATAATGAATATATCGGTTGCG 60
Db 1 ATGACGAATCTACCGATCAATAAATAATAGATGAAGTCAATAATGAATATATCGGTTGCG 60
QY 61 ATAAATAATCAGTAATGGATCTACGCTAGAGAAAGTATGTTGTTATTCATTAATGCT 120
Db 61 ATAAATAATCAGTAATGGATCTACGCTAGAGAAAGTATGTTGTTATTCATTAATGCT 120
QY 121 GGAGTAAACGATCCGACCGAGTCTGTTATTTACTCAGTTTATGATTCATTAATACCGAG 180
Db 121 GGAGTAAACGATCCGACCGAGTCTGTTATTTACTCAGTTTATGATTCATTAATACCGAG 180
QY 181 TATGAGTTAGGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCATCTATT 240
Db 181 TATGAGTTAGGTATGAAGAGCGCAATTCAGCTAGAAATGATTCATACATATTCATCTATT 240
QY 241 CATGATGACCTACCGAGCGGATGATTAATGATGATTCATGACGAGGAAATTAACAAATCAT 300
Db 241 CATGATGACCTACCGAGCGGATGATTAATGATGATTCATGACGAGGAAATTAACAAATCAT 300
QY 301 AAAGTATATGTTGAGTGCAGTCCGATATAGCAGGTGATGCTTTATTAACATAAGCATTT 360
Db 301 AAAGTATATGTTGAGTGCAGTCCGATATAGCAGGTGATGCTTTATTAACATAAGCATTT 360
QY 361 GAACCTATTTCAGGTATGATGATTAATGATGATTAATGATGATTAATGATGATTAATGATGAT 420
Db 361 GAACCTATTTCAGGTATGATGATTAATGATGATTAATGATGATTAATGATGATTAATGATGAT 420
QY 421 CTGCAATPAGCAAGTGGTCTGTTGGAATGTTGCGCGGTCAAAATGTTAGATGATCAAAAGC 480
Db 421 CTGCAATPAGCAAGTGGTCTGTTGGAATGTTGCGCGGTCAAAATGTTAGATGATCAAAAGC 480
QY 481 GAAGCCCAACCAATTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTT 540
Db 481 GAAGCCCAACCAATTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTTGAACTTT 540
QY 541 TTATTAACCTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATCACTAAA 600
Db 541 TTATTAACCTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATCACTAAA 600
QY 601 GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAGATGATTTA 660
Db 601 GAACATTTAGAAAGTTATAGTTATCATTTAGTATGATGTTCCAGATTAAGATGATTTA 660
QY 661 TTAGACTCTATGTTGATGAAGCAAGTTAGTAAAGTGGCGAGCATCTTCAAAAT 720
Db 661 TTAGACTCTATGTTGATGAAGCAAGTTAGTAAAGTGGCGAGCATCTTCAAAAT 720
QY 721 AATAAAGTACGTACGTGAGTTTATTTAGGAAAGATGGCGAGAGATAAATTTGACTTAT 780
Db 721 AATAAAGTACGTACGTGAGTTTATTTAGGAAAGATGGCGAGAGATAAATTTGACTTAT 780
QY 781 CATAGAGCGCAGCAGTGGATGAACATAACGAAATTTGATGAACATAACATAACAAACAC 840
Db 781 CATAGAGCGCAGCAGTGGATGAACATAACGAAATTTGATGAACATAACATAACAAACAC 840
QY 841 TTATTAGAAATCGTTGATTTA 861
Db 841 TTATTAGAAATCGTTGATTTA 861

RESULT 5
AX141501

LOCUS AX141501 783 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 223 from Patent WO0134809.
ACCESSION AX141501
VERSION AX141501.1 GI:14281622
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 783)
AUTHORS Kimmerly,W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 223 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
source
1. .783
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="synthetic nucleic acid sequence"
BASE COUNT 297 a 93 C 147 g 246 t
ORIGIN
Query Match 2.9%; Score 25; DB 6; Length 783;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 281 GAGGAAAATTACAAATCATAAAGT 305
Db 182 GAGGAAAATTACAAATCATAAAGT 206
RESULT 6
LOCUS AF270104 3444 bp DNA linear BCT 01-AUG-2000
DEFINITION Staphylococcus epidermidis strain SRI clone step.1049b08 genomic
sequence.
ACCESSION AF270104
VERSION AF270104.1 GI:9624010
KEYWORDS
SOURCE Staphylococcus epidermidis.
ORGANISM Staphylococcus epidermidis.
REFERENCE 1 (bases 1 to 3444)
AUTHORS Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3444)
AUTHORS Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source
1. .3444
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1049b08"
BASE COUNT 1365 a 451 c 571 g 1057 t
ORIGIN
Query Match 2.9%; Score 25; DB 1; Length 3444;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 281 GAGGAAAATTACAAATCATAAAGT 305
Db 182 GAGGAAAATTACAAATCATAAAGT 206
RESULT 6
LOCUS AF270104 3444 bp DNA linear BCT 01-AUG-2000
DEFINITION Staphylococcus epidermidis strain SRI clone step.1049b08 genomic
sequence.
ACCESSION AF270104
VERSION AF270104.1 GI:9624010
KEYWORDS
SOURCE Staphylococcus epidermidis.
ORGANISM Staphylococcus epidermidis.
REFERENCE 1 (bases 1 to 3444)
AUTHORS Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3444)
AUTHORS Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source
1. .3444
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1049b08"
BASE COUNT 1365 a 451 c 571 g 1057 t
ORIGIN
Query Match 2.9%; Score 25; DB 1; Length 3444;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 281 GAGGAAAATTACAAATCATAAAGT 305
Db 182 GAGGAAAATTACAAATCATAAAGT 206

QY 281 GAGGAAAATTACAAATCATAAAGT 305
Db 1046 GAGGAAAATTACAAATCATAAAGT 1070
RESULT 7
LOCUS AX145422 3444 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 4144 from Patent WO0134809.
ACCESSION AX145422
VERSION AX145422.1 GI:14283987
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 3444)
AUTHORS Kimmerly,W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 4144 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
source
1. .3444
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="synthetic nucleic acid sequence"
BASE COUNT 1365 a 451 c 571 g 1057 t
ORIGIN
Query Match 2.9%; Score 25; DB 6; Length 3444;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 281 GAGGAAAATTACAAATCATAAAGT 305
Db 1046 GAGGAAAATTACAAATCATAAAGT 1070
RESULT 8
LOCUS AF269958 3454 bp DNA linear BCT 01-AUG-2000
DEFINITION Staphylococcus epidermidis strain SRI clone step.1035g12 genomic
sequence.
ACCESSION AF269958
VERSION AF269958.1 GI:9623860
KEYWORDS
SOURCE Staphylococcus epidermidis.
ORGANISM Staphylococcus epidermidis.
REFERENCE 1 (bases 1 to 3454)
AUTHORS Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3454)
AUTHORS Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source
1. .3454
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1035g12"

BASE COUNT 1005 a 574 c 493 g 1382 t
ORIGIN

Query Match 2.9%; Score 25; DB 1; Length 3454;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GAGGAAATTAACAATCATAAAGT 305
|||||
Db 77 GAGGAAATTAACAATCATAAAGT 53

RESULT 9
AX145276/c
LOCUS AX145276 3454 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 3998 from Patent WO0134809.
ACCESSION AX145276
VERSION AX145276.1 GI:14283841
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
AUTHORS artificial sequences.
REFERENCE 1 (bases 1 to 3454)
TITLES Kimmerly, W.J.
JOURNAL Staphylococcus epidermidis nucleic acids and proteins
PATENT: WO 0134809-A 3998 17-MAY-2001;
GLAXO GROUP LIMITED (GB)

FEATURES
source
1. 3454
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"

BASE COUNT 1005 a 574 c 493 g 1382 t
ORIGIN

Query Match 2.9%; Score 25; DB 6; Length 3454;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GAGGAAATTAACAATCATAAAGT 305
|||||
Db 77 GAGGAAATTAACAATCATAAAGT 53

RESULT 10
AF269889/c
LOCUS AF269889 4045 bp DNA linear BCT 01-AUG-2000
DEFINITION Staphylococcus epidermidis strain SRI clone step.1028f08 genomic
sequence.
ACCESSION AF269889
VERSION AF269889.1 GI:9623789
KEYWORDS
SOURCE Staphylococcus epidermidis.
ORGANISM Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 4045)
TITLES Kimmerly, W.J., Taylor, J., David, N., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
Listembee, S., Ashanti, C., Altschuler, G., Mamo, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
JOURNAL Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
REFERENCE 2 (bases 1 to 4045)
TITLES Taylor, J., David, N., Nelsen, A.J., Godlevski, M.M.,
Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
Listembee, S., Ashanti, C., Altschuler, G., Mamo, L., Shepherd, N.S.,
Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
Furdon, P.J.
JOURNAL Direct Submission
COMMENT Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore.

Drive, Research Triangle Park, North Carolina 27709-3398, USA
Location/Qualifiers
1. 4045
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1028f08"

BASE COUNT 1249 a 691 c 536 g 1569 t
ORIGIN

Query Match 2.9%; Score 25; DB 1; Length 4045;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GAGGAAATTAACAATCATAAAGT 305
|||||
Db 2033 GAGGAAATTAACAATCATAAAGT 2009

RESULT 11
AX145207/c
LOCUS AX145207 4045 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 3929 from Patent WO0134809.
ACCESSION AX145207
VERSION AX145207.1 GI:14283772
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
AUTHORS artificial sequences.
REFERENCE 1 (bases 1 to 4045)
TITLES Kimmerly, W.J.
JOURNAL Staphylococcus epidermidis nucleic acids and proteins
PATENT: WO 0134809-A 3929 17-MAY-2001;
GLAXO GROUP LIMITED (GB)

FEATURES
source
1. 4045
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"

BASE COUNT 1249 a 691 c 536 g 1569 t
ORIGIN

Query Match 2.9%; Score 25; DB 6; Length 4045;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GAGGAAATTAACAATCATAAAGT 305
|||||
Db 2033 GAGGAAATTAACAATCATAAAGT 2009

RESULT 12
HS238J17
LOCUS HS238J17 111507 bp DNA linear PRI 23-NOV-1999
DEFINITION Human DNA sequence from PAC 238J17 on chromosome 6q22. Contains EST
and STS.
ACCESSION 298753
VERSION 298753.1 GI:2760552
KEYWORDS 6q22.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 111507)
TITLES Phillips, S.
JOURNAL Direct Submission
COMMENT Submitted (30-DEC-1997) Chromosome 6 Project Group
(http://www.sanger.ac.uk/HGP/Chr6/) Sanger Centre, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Jan 9, 1998 this sequence version replaced gi:2578075.
IMPORTANT: This sequence is the entire insert of clone 238J17.
During sequence assembly data is compared from overlapping clones.


```
/note="THEIB repeat: matches 1. .364 of consensus"
73212. .74406
/note="THEIB-INTERNAL repeat: matches 1. .1157 of
consensus"
74553. .75509
/note="MER11A repeat: matches 59. .1090 of consensus"
75510. .75941
/note="THEIB-INTERNAL repeat: matches 1153. .1580 of
consensus"
75942. .76305
/note="THEIB repeat: matches 1. .364 of consensus"
78184. .78607
/note="MSTA repeat: matches 424. .1 of consensus"
80178. .80727
/note="MLT2D repeat: matches 545. .1 of consensus"
82250. .82366
/note="MIR repeat: matches 157. .29 of consensus"
82329. .82669
/note="match: STS G05363"
85010. .85309
/note="AluSg repeat: matches 300. .1 of consensus"
86592. .86876
/note="AluY repeat: matches 293. .1 of consensus"
86901. .87116
/note="L1M3 repeat: matches 614. .399 of consensus"
86907. .87066
/note="L1MD2 repeat: matches 610. .457 of consensus"
87406. .92840
/note="L1 repeat: matches 11. .5390 of consensus"
92699. .93572
/note="L1P3 repeat: matches 1. .891 of consensus"
93896. .93921
/note="13 copies of 2 mer 100 % conserved"
94298. .94642

Query Match
Best Local Similarity 2.9% Score 25; DB 9; Length 111507;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 AACTGATGAAGTAAATAAAGTT 411
Db 46032 AACTGATGAAGTAAATAAAGTT 46056

RESULT 13
AC096106
LOCUS
DEFINITION Rattus norvegicus clone CH230-27L19, linear HTG 11-JUL-2002
*** 111 unordered pieces.
AC096106
VERSION AC096106.3 GI:21233786
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 227165)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbala,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowler,S., Brleva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burrett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.B., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
```

```
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Lucier,R., Luna,R., Ma,J.,
Lozaro,R.J., Lu,X., Lucier,A., Lucier,R., Liu,J., Liu,W., Louisedge,H.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,A., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,N., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivera,S., Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 227165)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 227165)
Worley,K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 29, 2002 this sequence version replaced gi:17943790.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GELF
Center clone name: CH230-27L19
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 113322 bases at least Q40
Consensus quality: 121538 bases at least Q30
Consensus quality: 128485 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 111 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1469: contig of 1469 bp in length
* 1470 1569: gap of unknown length
* 1570 2648: contig of 1079 bp in length
* 2649 2748: gap of unknown length
* 2749 4051: contig of 1303 bp in length
* 4052 4152: gap of unknown length
* 4152 5750: contig of 1599 bp in length
```

* 5751 5850: gap of unknown length
* 5851 6946: contig of 1096 bp in length
* 6947 7046: gap of unknown length
* 7047 8189: contig of 1143 bp in length
* 8190 8289: gap of unknown length
* 8290 9364: contig of 1075 bp in length
* 9365 9464: gap of unknown length
* 9465 10485: contig of 1021 bp in length
* 10486 10585: gap of unknown length
* 10586 12155: contig of 1570 bp in length
* 12156 12255: gap of unknown length
* 12256 13531: contig of 1276 bp in length
* 13532 13632: gap of unknown length
* 13633 15089: contig of 1458 bp in length
* 15090 15189: gap of unknown length
* 15190 16711: contig of 1522 bp in length
* 16712 17941: gap of unknown length
* 17942 18041: gap of unknown length
* 18042 19094: contig of 1052 bp in length
* 19094 19194: gap of unknown length
* 19194 20353: contig of 1160 bp in length
* 20354 20453: gap of unknown length
* 20454 21575: contig of 1122 bp in length
* 21576 21676: gap of unknown length
* 21677 23040: contig of 1365 bp in length
* 23041 23140: gap of unknown length
* 23141 24260: contig of 1120 bp in length
* 24261 24360: gap of unknown length
* 24361 25397: contig of 1036 bp in length
* 25397 25496: gap of unknown length
* 25497 27048: contig of 1552 bp in length
* 27049 27148: gap of unknown length
* 27149 28238: contig of 1090 bp in length
* 28239 28338: gap of unknown length
* 28339 30192: contig of 1854 bp in length
* 30193 30292: gap of unknown length
* 30293 31499: contig of 1207 bp in length
* 31500 31599: gap of unknown length
* 31600 33540: contig of 1941 bp in length
* 33541 33641: gap of unknown length
* 33641 35223: contig of 1583 bp in length
* 35224 35323: gap of unknown length
* 35324 36759: contig of 1436 bp in length
* 36760 36859: gap of unknown length
* 36860 38728: contig of 1869 bp in length
* 38729 38828: gap of unknown length
* 38829 40656: contig of 1828 bp in length
* 40657 40757: gap of unknown length
* 40758 42105: contig of 1349 bp in length
* 42106 42205: gap of unknown length
* 42206 43532: contig of 1327 bp in length
* 43533 43632: gap of unknown length
* 43633 45310: contig of 1678 bp in length
* 45311 45410: gap of unknown length
* 45411 46434: contig of 1024 bp in length
* 46435 46534: gap of unknown length
* 46535 47673: contig of 1139 bp in length
* 47674 47773: gap of unknown length
* 47774 49578: contig of 1805 bp in length
* 49579 49678: gap of unknown length
* 49679 51420: contig of 1742 bp in length
* 51421 51520: gap of unknown length
* 51521 52838: contig of 1318 bp in length
* 52839 52938: gap of unknown length
* 52939 54588: contig of 1650 bp in length
* 54589 54688: gap of unknown length
* 54689 55991: contig of 1303 bp in length
* 55992 56091: gap of unknown length
* 56092 57458: contig of 1367 bp in length
* 57459 57558: gap of unknown length
* 57559 58637: contig of 1079 bp in length
* 58638 58737: gap of unknown length

* 58738 60572: contig of 1835 bp in length
* 60573 60672: gap of unknown length
* 60673 61999: contig of 1327 bp in length
* 62000 62099: gap of unknown length
* 62100 63876: contig of 1777 bp in length
* 63877 63976: gap of unknown length
* 63977 65299: contig of 1323 bp in length
* 65300 65399: gap of unknown length
* 65400 66581: contig of 1182 bp in length
* 66582 66681: gap of unknown length
* 66682 67940: contig of 1259 bp in length
* 67941 68040: gap of unknown length
* 68041 69984: contig of 1944 bp in length
* 69985 70084: gap of unknown length
* 70085 71379: contig of 1295 bp in length
* 71380 71479: gap of unknown length
* 71480 73359: contig of 1880 bp in length
* 73360 73459: gap of unknown length
* 73460 74946: contig of 1487 bp in length
* 74947 75046: gap of unknown length
* 75047 76380: contig of 1334 bp in length
* 76381 76480: gap of unknown length
* 76481 78494: contig of 2014 bp in length
* 78495 78594: gap of unknown length
* 78595 79825: contig of 1231 bp in length
* 79826 79925: gap of unknown length
* 79926 81163: contig of 1238 bp in length

Query Match 2.6%; Score 22; DB 2; Length 227165;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 AACTGATGAAGTAAATAAATAAA 408
|||||
Db 157861 AACTGATGAAGTAAATAAATAAA 157882

RESULT 14

CPU30821/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

JOURNAL

FEATURES

source

gene

CDS

CPU30821 135599 bp DNA circular PLN 13-NOV-1995
Cyanophora paradoxa cyanelle, complete genome.

U30821
U30821.1 GI:1016083

Cyanophora paradoxa.
Cyanelle Cyanophora paradoxa
Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.

1 (bases 1 to 135599)
Stirewalt, V.L., Michalowski, C.B., Luffelhardt, W., Bohnert, H.J. and
Bryant, D.A.

Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa

Unpublished
2 (bases 1 to 135599)
Bryant, D.A.

Direct Submission
Submitted (01-JUL-1995) Donald A. Bryant, Biochemistry and
Molecular Biology, The Pennsylvania State University, S-234 Frear
Bldg., University Park, PA 16802, USA

Location/Qualifiers
1. .135599
/organism="Cyanophora paradoxa"
/organelle="plastid:cyanelle"
/strain="Pringsheim L8555"
/db_xref="taxon:2762"
/gene="ycf27"
complement(join(134890..135599,1..10))
complement(join(134890..135599,1..10))
/function="transcriptional regulator of the OmpR family;
target unknown"
/codon_start=1
/protein_id="AA81319.1"
/db_xref="GI:1016232"

```

/translacion="MEINKKILVWDEISIRRIETRLSIMGVYVTAADGERALII
FOLZHPVLVDVMPKLDGIGVQCUELRKESDIPILMLTALGDVADRITGLGADDI
VWPKFSGLELARISRLRRTENPTLSTSTNGENLQIGFLKIDINKRQFKNGERIR
LTKGFEILLLELISKMGEPFSRAQILQEVNGYTPERHIDTRVVDVHISRLRSKLEENP
SNPLIITARGIGLYFQNTSN"
1. .11285
/feature="IR-A"
/rpt_type="inverted"
294. .1788
/gene="rrsA-A"
294. .1788
/gene="rrsA-A"
/product="16S ribosomal RNA"
1889. .1962
/gene="trnI-A"
1889. .1962
/gene="trnI-A"
/product="tRNA-Ile"
/feature="anticodon (GAU)"
1966. .2038
/gene="trnA-A"
1966. .2038
/gene="trnA-A"
/product="tRNA-Ala"
/feature="anticodon (UGC)"
2076. .5001
/gene="rrfA-A"
2076. .5001
/gene="rrfA-A"
/product="23S ribosomal RNA"
5102. .5219
/gene="rrfA-A"
5102. .5219
/gene="rrfA-A"
/product="5S ribosomal RNA"
5727. .7598
/gene="dnaK-A"
5727. .7598
/gene="dnaK-A"
/function="molecular chaperone; heat-shock protein;
protein folding"
/codon_start=1
/product="DnaK"
/protein_id="AAA81171.1"
/db_xref="GI:1016084"
/translacion="MEVHKMGKIVGIDLTGNSVAVMEGKPVITNAEGFRTPSV
VAYTKGDLVGOIAKQAVLNPCTFYSVKRFGRKDFVTEAKOVYPKYVTQTVY
RIECPALQTSPTPEISAOVLKLAEDASKYIGEPVTOAVITVPAYFNDQSQAQKDA
GKIAGLEVLRINEPTAASLAYGLDKKNEIILVFDLGGGTFDVSILEVGCVEVLS
TSGTHLGGDDFDKLLVDFFLEEFQKLENINREDQALQRLTEAAEKAKTELSVTO
TEITLPFLSTPSPGPKHFDFTLNRAKFEELGSLTIDRCRIPVETALKDAKTKEDIDE
VVLVGGSTRIPAVRELQKVLDPKQSVNPDEVAVGAAGVQAGVLAGVKDILVLDY
TPLSIGVETLGGVMSKIVRNTTIPQKSELFSTATDQTNVEIHVLQGERGFARDNK
SLGTRFDLGTAPAPRGVPOIEVTFDIDANGILSVTAKDKGKREKSIITGASFLSDQ
EYERVOEAKENQAQDKENQKIELRQAESLILNQAEKQLKEFKLEPNVQAPISQTIQ
QIQTALKEDNLEQVQSCQELQNSLLEIGKQVYSKITEITIPKNEEDVIDASFSEK"
7694. .7764
/gene="trnC-A"
7694. .7764
/gene="trnC-A"
/product="tRNA-Cys"
/feature="anticodon (GCA)"
complement(8038. .9663)
/gene="groEL-A"
complement(8038. .9663)
/gene="groEL-A"
/function="molecular chaperone; heat-shock protein;
protein folding"
/codon_start=1
/product="GroEL"
/protein_id="AAA81172.1"
/db_xref="GI:1016085"
/translacion="MAQILYHESARRALEKMGDILAEAVNTLGPGRNVVLEKKFG

```

```

APQIVNDGVTTAKIEIELEDHIENTGVALLIRQAAASKTNDVAGDGTATTATVLAYAMKEG
MRNVAAGNPAIARKIDRATQKVVETIAAHAKPVEDSRATQVASISAGNDEVGKM
IADAIEKVGKGVISLEEGSTTELEITEGMRFDKGYISPYFVDPMEVIOQDAM
FLIDKKITVLQDVLVPLEQVARARKPLFIIAEDIEKEALATLVVNLKRLNLTSAVK
APGPGDRKAMQDIAVLTNTQVTSQTLGLDSDIKLDVLGKAKQVNTIKDQYTVIAE
GNEEVSARCEQIRROTEETSAYEKEKLOERLAKAGGVAVIKVGAATEMEKDKKL
RLEDAINNTKAAVEEGIVPGGTTLAHLAIELENWQVSNLEHEELIGMIVARSILTAP
LKRIVENNGAGVIAQVQKDPENIGNAANAQFVDMFEAGIVDPKAVTRSLQNA
SIAGMVLTTCEIVVDKAEQQAASVNRNSPDY"
complement(9727. .10038)
/gene="groES-A"
complement(9727. .10038)
/gene="groES-A"
/function="molecular chaperone; chaperonin; heat-shock
protein; protein folding"
/codon_start=1
/product="GroES"
/protein_id="AAA81173.1"
/db_xref="GI:1016086"
/translacion="MAVTLNVKTVRPLGERVLVVKVQSSEKTAGGILLPDTVKEKPO
IGETIABGPGRRNDGSGFPLEVTNVSKVLSKYAGTDIKLENEYVLLSEKDLIAII
A"
10419. .11003
/gene="clpP"
10419. .11003
/gene="clpP"
/function="ATP-dependent protease subunit"
/codon_start=1
/product="ClpP protease subunit"
/protein_id="AAA81174.1"
/db_xref="GI:1016087"
/translacion="MTILIQYTDNKKLENNISKLLQNRILFQSVTDETCNSI
VGQLLYENEDSTKDIRLFINPGSVTAGLSVDYTIQNLSDVSTICFGLAASMGAV
LLAAGVENKRFAPASIRIMHQPLSKVEAPWSHLDIQIRNGAYFNKLLNLSFPTKQ
ELKQIETDTERDFLSAPEAKQYGIIDHIFINN"
11515. .11586
/gene="trnK"
11515. .11586
/gene="trnK"
/product="tRNA-Lys"
/feature="anticodon (UUU)"
complement(11591. .12109)
/gene="ycf37"
complement(11591. .12109)
/gene="ycf37"
/codon_start=1
/protein_id="AAA81175.1"
/db_xref="GI:1016088"
/translacion="MDNLFSEYIIFSLIVLALISFLIQIKSRFLDLLETRLRKK
IYQEKATEKYNLNASIYIKRKLKYSQAKELKLENNKLDKQALINSLGYTCSA
QEQYELAIEYKKALFYIPDFILARINLARIFELKLNKTEAINMYQEVLIQPKNOLA
KRKLSIITENLIL"
complement(12224. .12457)
/gene="ycf34"
complement(12224. .12457)
/gene="ycf34"
/codon_start=1
/protein_id="AAA81176.1"
/db_xref="GI:1016089"
/translacion="MCICVNDLLVFCSLYQTIETLHKESLSFELDQIVEPEPRKSN
IVNINSSENSIKHEWDVSCSFVNLKSNVS"
complement(join(12625. .12674,12907. .12941))
/gene="trnL-A"
complement(join(12625. .12674,12907. .12941))
/gene="trnL-A"
/product="tRNA-Leu"
/feature="anticodon (UAA)"
complement(12675. .12906)
/gene="trnL-A"
/feature="Type-I intron; found in many cyanobacteria and some
eubacteria"
complement(13099. .13171)
/gene="trnE"

```

```
trna      complement(13099..13171)
/gene="trnE"
gene      /note="anticodon (UUC)"
13520..13831
/gene="acpp"
CDS       13520..13831
/gene="acpp"
/function="fatty acid biosynthesis"
/codon_start=1
/product="acyl carrier protein"
/db_xref="GI:1016090"
/translation="MSLSQAILEQDKNIVNQDILCEQISIEKTEINLQDIYED
LKIDSULVEIIKQIEEKDKIDSKILYMTLEFIDFTLQTIYMKHGFYLNKQ
F"
gene      13893..14456
/gene="ycf4"
CDS       13893..14456
/gene="ycf4"
/function="possible in regulating PS I:PS II ratio"

Query Match      2.4%; Score 21; DB 8; Length 135599;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      328 TTAGCAGGTGATGCTTTATTA 348
          |||||
Db      129578 TTAGCAGGTGATGCTTTATTA 129558

RESULT 15
AF236874      157270 bp DNA linear PRI 02-MAR-2000
LOCUS      Homo sapiens chromosome 21 map 21q21, BAC B242B02, complete
DEFINITION      sequence.
ACCESSION      AF236874
VERSION      AF236874.1 GI:7141351
KEYWORDS      HTG.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 157270)
Bleichschmidt,K., Riesselmann,L., Wehrmeyer,S., Baumgart,C.,
Menzel,U., Dette,M., Jahn,N., Schilhabel,M., Yaspo,M.-L. and
Rosenthal,A.
TITLE      Direct Submission
JOURNAL      Submitted (18-FEB-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
source
1..157270
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q21"
/clone="BAC B242B02"
repeat_region complement(1..134)
repeat_region /rpt_family="MLT1F"
repeat_region complement(19..429)
repeat_region /rpt_family="MLT1E"
repeat_region 2089..2400
repeat_region /rpt_family="AluSq"
repeat_region 2700..2802
repeat_region /rpt_family="MER81"
exon 2730..2840
/note="GRAIL"
evidence-not_experimental
repeat_region complement(2943..3223)
repeat_region /rpt_family="MLT1E"
repeat_region 3798..4553
repeat_region /rpt_family="L1MCB_Send"
repeat_region complement(3824..4287)
repeat_region /rpt_family="MLT1D"

4542..4639
/note="GRAIL"
evidence-not_experimental
repeat_region complement(4817..5043)
/rpt_family="LOR1a"
repeat_region complement(5041..5106)
/rpt_family="LOR1a"
repeat_region complement(5112..5402)
/rpt_family="AluSx"
repeat_region 5648..6167
/rpt_family="L1MCB_Send"
repeat_region 6180..6502
/rpt_family="L1MEC_Send"
repeat_region complement(6613..7086)
/rpt_family="L1PAL5"
repeat_region complement(7161..7578)
/rpt_family="L1PAL4"
repeat_region complement(7436..11435)
/rpt_family="L1"
repeat_region 7861..7884
/note="(GGA)n"
/rpt_family="Simple repeat"
complement(8364..8513)
/note="GRAIL"
evidence-not_experimental
repeat_region complement(11225..11846)
/rpt_family="L1PAL2_Send"
repeat_region complement(11516..12125)
/rpt_family="L1PAL3_Send"
complement(12550..12611)
/note="MZF"
evidence-not_experimental
exon complement(12810..12873)
/note="MZF"
evidence-not_experimental
exon 13577..13759
/note="MZF"
evidence-not_experimental
repeat_region 14607..14633
/note="AT rich"
repeat_region /rpt_family="Low complexity"
complement(15088..15448)
/rpt_family="THE1B"
exon 15178..15330
/note="GRAIL"
evidence-not_experimental
repeat_region complement(15947..15997)
/rpt_family="L1MA9"
repeat_region 16433..16457
/note="AT rich"
/rpt_family="L2a"
repeat_region complement(17199..17507)
/rpt_family="L2a"
exon 17452..17562
/note="GENSCAN"
repeat_region complement(17506..18111)
/rpt_family="L2"
exon 17795..18032
/note="GRAIL"
evidence-not_experimental
repeat_region 18238..18500
/rpt_family="MLT1H"
repeat_region complement(18825..18851)
/note="AT rich"
repeat_region /rpt_family="Low complexity"
complement(18878..18898)
/rpt_family="AT rich"
repeat_region complement(18938..18977)
/note="AT rich"
repeat_region /rpt_family="Low complexity"
complement(19054..19199)
/note="GRAIL"
exon
```

```
exon      /evidence-not_experimental
19154..19261
/note="GENSCAN"
complement(20188..20326)
/note="GRAIL"
/evidence-not_experimental
repeat_region complement(20589..20863)
/rpt_family="MERSA"
repeat_region complement(21353..21379)
/note="AT-rich"
exon      /rpt_family="Low_complexity"
21378..21447
/note="GRAIL"
/evidence-not_experimental
repeat_region complement(21509..21667)
/rpt_family="AluSg1"
exon      complement(21638..21785)
/note="MZER"
/evidence-not_experimental
repeat_region 22368..22902
/rpt_family="MLTIE"
repeat_region 24714..25090
/rpt_family="MERS7B"
repeat_region complement(25217..25525)
/rpt_family="AluSc"
exon      25801..25905
/note="MZER"
/evidence-not_experimental
repeat_region 25807..26115
/rpt_family="AluY"
exon      complement(26760..26806)
/note="MZER"
repeat_region /evidence-not_experimental
complement(27112..27454)
/note="(GA)n"
repeat_region /rpt_family="Simple_repeat"
complement(27132..27443)
/rpt_family="AluSp"
exon      28969..29151
/note="GRAIL"
/evidence-not_experimental
repeat_region complement(29179..29264)
/rpt_family="L1PA8"
repeat_region 29727..30189
/rpt_family="MER31A"
repeat_region 30395..30522
/rpt_family="MIR"
exon      30759..30834
/note="GRAIL"
/evidence-not_experimental
repeat_region 31435..31607
/note="AT-rich"
repeat_region /rpt_family="Low_complexity"
complement(31610..31678)
/note="(TA)n"
exon      /rpt_family="Simple_repeat"
complement(31923..31969)
/note="XPOUND"
exon      /evidence-not_experimental
complement(33187..33252)
/note="GRAIL"
repeat_region /evidence-not_experimental
complement(33838..34147)
/rpt_family="AluSc"
repeat_region complement(34805..34853)
/rpt_family="AluSg"
repeat_region 35084..35200
/rpt_family="L1MC5"
repeat_region complement(35212..35730)
/rpt_family="L1"
exon      complement(35238..35438)
/note="GRAIL"
/evidence-not_experimental
```

```
exon      complement(35238..35315)
/note="MZER"
/evidence-not_experimental
exon      complement(35326..35375)
/note="XPOUND"
/evidence-not_experimental
exon      complement(35438..35682)
/note="GENSCAN"
exon      complement(35438..35682)
/note="GRAIL"
repeat_region /evidence-not_experimental
35725..36564
/rpt_family="L1"
```

Query Match 2.4%; Score 21; DB 9; Length 157270;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 346 TTAAGTAAAGCATTTGAACCTT 366
DB 81200 TTAAGTAAAGCATTTGAACCTT 81220

Search completed: May 30, 2003, 19:07:51
Job time : 2487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 17:23:55 ; Search time 280 Seconds
(without alignments)
6924.895 Million cell updates/sec

Title: US-09-925-637-63

Perfect score: 861

Sequence: 1 atagcaattaccagtgaa.....tattagaatcgttgattta 861

Scoring table: OIIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SID22/qcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/qcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/qcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/qcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/qcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/qcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/qcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/qcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/qcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/qcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/qcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/qcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/qcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/qcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/qcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/qcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/qcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/qcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/qcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/qcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/qcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/qcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/qcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/qcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	861	100.0	861	22	AAS00821		S. aureus HGS072 e
2	807	93.7	1893	18	AAV74466		Staphylococcus aur
3	606	70.4	882	21	AA92031		Staphylococcus aur
4	489	56.8	882	23	AA54848		Staphylococcus aur
5	474	55.1	864	23	AA551602		Staphylococcus aur
6	413	48.0	413	23	AA550245		Staphylococcus aur
7	337	39.1	337	23	AA550698		Staphylococcus aur
8	25	2.9	783	22	AAH52415		S. epidermidis ope
9	25	2.9	909	24	ABN90871		Staphylococcus epi

10	25	2.9	3444	22	AAH54780	S. epidermidis gen
11	25	2.9	3454	22	AAH54634	S. epidermidis gen
12	25	2.9	4045	22	AAH54565	S. epidermidis gen
13	20	2.3	2365589	24	ABA90521	Genomic sequence o
14	19	2.2	151	22	ABA74941	Human foetal liver
15	19	2.2	151	22	ABA39630	Probe #18096 for q
16	19	2.2	151	22	ABA33438	Human brain expres
17	19	2.2	151	22	AAK49586	Human bone marrow
18	19	2.2	151	22	AAI26694	Probe #16627 for q
19	19	2.2	151	22	AAI55466	Probe #24152 used
20	19	2.2	151	24	ABS23069	Human genome-deriv
21	19	2.2	600	22	ABA62413	Human foetal liver
22	19	2.2	600	22	ABA29746	Probe #8212 for q
23	19	2.2	600	22	AAK10751	Human brain expres
24	19	2.2	600	22	AAK36630	Human bone marrow
25	19	2.2	600	22	AAI17482	Probe #7415 for q
26	19	2.2	600	22	AAI42391	Probe #11077 used
27	19	2.2	600	24	ABS10626	Human genome-deriv
28	19	2.2	809	24	ABQ33820	Oligonucleotide fo
29	19	2.2	809	24	ABQ33821	Oligonucleotide fo
30	19	2.2	873	24	ABN66184	Streptococcus poly
31	19	2.2	6072	23	ABL07772	Drosophila melanog
32	19	2.2	640681	24	ABA92787	Buchnera sp. genom
33	19	2.2	2155561	24	ABN71527	Streptococcus poly
34	18	2.1	60	19	AAV45107	SDF-1-alpha primer
35	18	2.1	118	22	AAS30093	Human lung antigen
36	18	2.1	164	21	AAC07133	Human secreted pro
37	18	2.1	322	20	AAV87896	EST clone FG99. H
38	18	2.1	350	23	ABV50402	Human prostate exp
39	18	2.1	491	24	ABN61915	Human cancer relat
40	18	2.1	588	22	AAH12725	Human cDNA clone (
41	18	2.1	721	22	AAS27074	cDNA encoding nove
42	18	2.1	721	23	ABK43767	DNA encoding novel
43	18	2.1	919	21	AAC56354	Pinus radiata tran
44	18	2.1	1301	12	AAQ14532	T cell protein 19.
45	18	2.1	1301	15	AAQ63615	T-cell transmembra

ALIGNMENTS

RESULT 1
AAS00821
ID AAS00821 standard; DNA; 861 BP.
AC AAS00821;
XX
XX
DT 04-JUL-2001 (first entry)
DE S. aureus HGS072 encoding Farnesyl diphosphatesynthase, Ispa.
XX
KW Farnesyl diphosphatesynthase; Ispa; immunogen; vaccine; antibody;
KW wound infection; cellulitis; burn infection; eyelid infection;
KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
KW HGS072; qs.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..861
FT FT /*tag- a
FT FT /product- "IspsA"
FT FT /partial
FT FT /note- "No stop codon"

WO200116292-A2.
08-MAR-2001.
31-AUG-2000; 2000WO-US23773.

Db	481	GAAGGCCAACCAATTGATCTTTGAAACTTTGGAAATGATACACAAAAACAGGAGCA	540
Qy	541	TTATTAACCTTTTGCGGTTATAGTGCAGCAGATATCGCTAATGTGATGATACAACTAA	600
Db	541	TTATTAACCTTTTGCGGTTATAGTGCAGCAGATATCGCTAATGTGATGATACAACTAA	600
Qy	601	GAACATTTAGAAAGTTATAGTTATCATCTTTAGTATGATCTTCCAGATTTAAAGATGATTTA	660
Db	601	GAACATTTAGAAAGTTATAGTTATCATCTTTAGTATGATCTTCCAGATTTAAAGATGATTTA	660
Qy	661	TTAGACTCTATGGTGATGAAGCAAAAGTTAGTAAAAAGTGGCGAGCGATCTTTGAAAAT	720
Db	661	TTAGACTCTATGGTGATGAAGCAAAAGTTAGTAAAAAGTGGCGAGCGATCTTTGAAAAT	720
Qy	721	AATAAAGTACGTACGTGAGTTTATTAGGGAAGATGGCGCAGAGATTAATTGACTTAT	780
Db	721	AATAAAGTACGTACGTGAGTTTATTAGGGAAGATGGCGCAGAGATTAATTGACTTAT	780
Qy	781	CATAGAGACGCAGCAGTGGATCAACTAAACGCAAAATTGATGAACAAATTCATACAAAAACAC	840
Db	781	CATAGAGACGCAGCAGTGGATCAACTAAACGCAAAATTGATGAACAAATTCATACAAAAACAC	840
Qy	841	TTATTAGAAATCGTTGATTTA 861	
Db	841	TTATTAGAAATCGTTGATTTA 861	
RESULT 2			
AAV74466/c			
ID	AAV74466 standard; DNA; 1893 BP.		
XX			
AC	AAV74466;		
XX			
DT	16-MAR-1999 (first entry)		
XX			
DE	Staphylococcus aureus contig SEQ ID #155.		
XX			
KW	Computer readable medium; vaccine; S.aureus infection; immunodetection;		
KW	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;		
KW	skin infection; surgical wound infection; scalded skin syndrome;		
KW	toxic shock syndrome; ds.		
XX			
OS	Staphylococcus aureus.		
XX			
FI	Key	Location/Qualifiers	
FT	misc_feature	1261..1320	
FT		/*tag= a	
FT		/note= "these bases represent a line of missing text in	
FT		the sequence listing in the specification. They	
FT		are included to maintain the nucleotide numbering	
FT		given in the specification for this DNA sequence"	
XX			
PN	EP786519-A2.		
XX			
PD	30-JUL-1997.		
XX			
PF	07-JAN-1997; 97EP-0100117.		
XX			
PR	05-JAN-1996; 96US-0009861.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;		
FI	Rosen CA;		
XX			
DR	WPI; 1997-374922/35.		
XX			
PT	Polynucleotide(s) and proteins derived from Staphylococcus aureus		
PT	stored on computer readable medium and used in the production of		
FT	anti-S.aureus vaccines		
XX			
PS	Claim 1; Page 784-785; 327lpp; English.		
XX			

CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.

XX Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

Query Match 93.7%; Score 807; DB 18; Length 1893;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	55	GTTCGGATAAATAATCAGTAATCGATGACTCAGCTAGAGAAATATGTTGTTATTCATTA	114
Db	1260	GTTCGGATAAATAATCAGTAATCGATGACTCAGCTAGAGAAATATGTTGTTATTCATTA	1201
Qy	115	AATCCTGGAGTAAACCCATCCGACCACTGCTGTTATTTACTCCTTAGATTCACTAAAT	174
Db	1200	AATCCTGGAGTAAACCCATCCGACCACTGCTGTTATTTACTCCTTAGATTCACTAAAT	1141
Qy	175	ACCGAGTATGAGTTAGTATGAGAGCGGCAATTCACACTAGAAATGATTACATATTTCA	234
Db	1140	ACCGAGTATGAGTTAGTATGAGAGCGGCAATTCACACTAGAAATGATTACATATTTCA	1081
Qy	235	CTTATTATGATGACCTACCGCATGATGATGATGATGATGATGATGATGATGATGATG	294
Db	1080	CTTATTATGATGACCTACCGCATGATGATGATGATGATGATGATGATGATGATGATG	1021
Qy	295	AATCATAAAGTATATGTTGAGTGACCTGCGATATTACAGGTGATGCTTTATTAACATAA	354
Db	1020	AATCATAAAGTATATGTTGAGTGACCTGCGATATTACAGGTGATGCTTTATTAACATAA	961
Qy	355	GCATTTGAACCTATTTCAGTGATGATGATGATGATGATGATGATGATGATGATGATG	414
Db	960	GCATTTGAACCTATTTCAGTGATGATGATGATGATGATGATGATGATGATGATGATG	901
Db	415	CAACGGCTGTCATAGCAAGTGGTCATGTTGGATGTTGCGGCTGCAAACTTAGATATG	474
Db	900	CAACGGCTGTCATAGCAAGTGGTCATGTTGGATGTTGCGGCTGCAAACTTAGATATG	841
Qy	475	CAAAAGCAAGGCCAACCAATTTGATCTTTGAAATTTGGAATGATACACAAAACAAAACA	534
Db	840	CAAAAGCAAGGCCAACCAATTTGATCTTTGAAATTTGGAATGATACACAAAACAAAACA	781
Qy	535	GGAGCAATTAATTAATTTTTCGGGTTATGAGTGCAGCAGATATGCTAATGCTGATGATAC	594
Db	780	GGAGCAATTAATTAATTTTTCGGGTTATGAGTGCAGCAGATATGCTAATGCTGATGATAC	721
Qy	595	ACTAAAGCAACATTTAGAAAGTTATGATATATCATTTAGTATGATGTTCCAGATTAAGAT	654
Db	720	ACTAAAGCAACATTTAGAAAGTTATGATATATCATTTAGTATGATGTTCCAGATTAAGAT	661
Qy	655	GATTTATTAGACTGCTATGTTGATGAAGCAAGTTAGTAAAAAGTGGCGAGCTCTT	714
Db	660	GATTTATTAGACTGCTATGTTGATGAAGCAAGTTAGTAAAAAGTGGCGAGCTCTT	601
Qy	715	GAAATATTAAGTACGTACGTGATTTATTAGGAAAGATGCGCGAGAGATAAATG	774
Db	600	GAAATATTAAGTACGTACGTGATTTATTAGGAAAGATGCGCGAGAGATAAATG	541

Qy	775	ACTTATCATAGACGACGACGAGTGGATGAACCTAACGCAAAATTTGATGAACAATTCATATACA	834
Db	540	ACTTATCATAGACGACGACGAGTGGATGAACCTAACGCAAAATTTGATGAACAATTCATATACA	481
Qy	835	AAACACTTATTAGAAATCGTTGATTTA	861
Db	480	AAACACTTATTAGAAATCGTTGATTTA	454

RESULT 3

ID	AAA92031	standard; DNA; 882 BP.
XX	AAA92031;	
AC	AAA92031;	
XX	12-JAN-2001	(first entry)
DT		
XX	Staphylococcus aureus ispa coding sequence.	
DE		
XX	IsPa; bacterial disease; respiratory tract infection;	
KW	gastrointestinal infection; cardiac infection; Helicobacter pylori;	
KW	stomach cancer; stomach ulcer; gastritis; ds.	
XX	Staphylococcus aureus.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..882
FT		/*tag= a
FT		/product= "IsPa"
XX	US6107058-A.	
PN		
XX	22-AUG-2000.	
PD		
XX	26-MAR-1999;	99US-0276873.
PF		
XX	26-MAR-1999;	99US-0276873.
PR		
XX	(SMIK) SMITHKLINE BEECHAM CORP.	
PA		
XX	Gwynn M, Wilding EI;	
PI		
XX	WPI; 2000-578535/54.	
DR	P-PSDB; AAB23333.	
DR		
XX		

Novel farnesyl diphosphate synthase polynucleotide from staphylococcus aureus useful for diagnosis and treatment of bacterial infections and as hybridization probe for isolating genomic clones -

Claim 1; column 1-4; 15pp; English.

The present sequence is the coding sequence for the Staphylococcus aureus IsPa protein. This gene and the protein it encodes can be used in many research assays, as well as treatments for bacterial diseases such as infections of the respiratory tract (including otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis, empyema and lung abscesses), cardiac infections such as infective endocarditis, gastrointestinal infections including secretory diarrhoea, splenic abscesses and retrofemoral abscesses, CNS infections such as cerebral abscesses, eye infections (including blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis and dacryocystitis), kidney and urinary tract infections such as epididymitis, intrarenal and perinephric abscesses and toxic shock syndrome, skin diseases (including impetigo, folliculitis, cutaneous abscesses, wound infection and bacterial myositis), bone and joint infections such as septic arthritis and osteomyelitis, septic thrombophlebitis, food poisoning and scalded skin syndrome. In addition, they can be used to treat diseases caused by Helicobacter pylori, including stomach cancer, stomach ulcers and gastritis.

Sequence 882 BP; 328 A; 118 C; 186 G; 250 T; 0 other;

Query Match 70.4%; Score 606; DB 21; Length 882;

Best Local Similarity 99.4%; Pred. No. 4.6e-291;
Matches 856; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGACGAATCTACCGATGAATAAATAAGTAGAAGTCAATTAATATCGGTGG 60
DB 1 ATGACGAATCTACCGATGAATAAATAAGTAGAAGTCAATTAATATCGGTGG 60

QY 61 ATAAATAAATACAGTAATGGATCTACAGTAGAAGAAAGTATGTATTCAATTAATGCT 120
DB 61 ATAAATAAATACAGTAATGGATCTACAGTAGAAGAAAGTATGTATTCAATTAATGCT 120

QY 121 GGAGGTAACGCATCCGACAGTCTGTATTACTACCTTTAGATTCACATAATCCGAG 180
DB 121 GGAGGTAACGCATCCGACAGTCTGTATTACTACCTTTAGATTCACATAATCCGAG 180

QY 181 TATGAGTTAGGTATGAAGAGCGCAATTCGACTAGAAATGATTCACATATTCACATTAT 240
DB 181 TATGAGTTAGGTATGAAGAGCGCAATTCGACTAGAAATGATTCACATATTCACATTAT 240

QY 241 CATGATGACCTACCGATGGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 CATGATGACCTACCGATGGATGATGATGATGATGATGATGATGATGATGATGATGAT 300

QY 301 AAGATATATGAGTGGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 AAGATATATGAGTGGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 360

QY 361 GAATTTATTTCAAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 GAATTTATTTCAAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

QY 421 CTGTCATAGCAAGTGGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 CTGTCATAGCAAGTGGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

QY 481 GAAGGCCAACCAATTCATCTGAACTTTGGAATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GAAGGCCAACCAATTCATCTGAACTTTGGAATGATGATGATGATGATGATGATGATGATGAT 540

QY 541 TTATTAACCTTTGCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 TTATTAACCTTTGCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

QY 601 GAACATTTAGAAAGTTATAGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 GAACATTTAGAAAGTTATAGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 660

QY 661 TTAGACTGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 TTAGACTGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

QY 721 AATAAAGTACGTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 AATAAAGTACGTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 781 CATAGAGCGCAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CATAGAGCGCAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

QY 841 TTATTAGAAATCGTTGATTTA 861
DB 841 TTATTAGAAATCGTTGATTTA 861

RESULT 4
AAS54848
ID AAS54848 standard; DNA; 882 BP.
XX
AC AAS54848;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #1160.

XX Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX Staphylococcus aureus.
OS
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-369308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR
DR R-PSDB; AAU36989.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Claim 27; Seq ID No 8485; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 882 BP; 330 A; 116 C; 185 G; 251 T; 0 other;

Query Match 56.8%; Score 489; DB 23; Length 882;
Best Local Similarity 99.8%; Pred. No. 6.5e-233;
Matches 539; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACGAATCTACCGATGAATAAATAAGTAGAAGTCAATTAATATCGGTGG 60
DB 1 ATGACGAATCTACCGATGAATAAATAAGTAGAAGTCAATTAATATCGGTGG 60

QY 61 ATAAATAAATACAGTAATGGATCTACAGTAGAAGAAAGTATGTATTCAATTAATGCT 120
DB 61 ATAAATAAATACAGTAATGGATCTACAGTAGAAGAAAGTATGTATTCAATTAATGCT 120

QY 121 GGAGGTAACGCATCCGACAGTCTGTATTACTACCTTTAGATTCACATAATCCGAG 180
DB 121 GGAGGTAACGCATCCGACAGTCTGTATTACTACCTTTAGATTCACATAATCCGAG 180

QY 181 TATGAGTTAGGTATGAAGAGCGCAATTCGACTAGAAATGATTCACATATTCACATTAT 240
DB 181 TATGAGTTAGGTATGAAGAGCGCAATTCGACTAGAAATGATTCACATATTCACATTAT 240

Db 181 TATGAGTTAGTATCAAGAGCCCAATTGCACTAGAAAATGATTCATACATATTCCTACTTAT 240
QY 241 CATGATGACCTACCAAGCGATGATATGATGATTCGACGAGGAAAATTAACAATCAT 300
Db 241 CATGATGACCTACCAAGCGATGATATGATGATTCGACGAGGAAAATTAACAATCAT 300
QY 301 AAAGTATATGTTGAGTGCAGTCGCGATATAGCAGGTGATGCTTTTATTAACAAACATTT 360
Db 301 AAAGTATATGTTGAGTGCAGTCGCGATATAGCAGGTGATGCTTTTATTAACAAACATTT 360
QY 361 GAACCTATTTCGAAGTGATGATGATTAAGTGAAGTAAATAAATAAAGTTCTCAACAGG 420
Db 361 GAACCTATTTCGAAGTGATGATGATTAAGTGAAGTAAATAAATAAAGTTCTCAACAGG 420
QY 421 CTGTCAATAGCAAGTGGTCATGTTGGAAATGTCGCGGTCAAATGTTAGATATGCAAGC 480
Db 421 CTGTCAATAGCAAGTGGTCATGTTGGAAATGTCGCGGTCAAATGTTAGATATGCAAGC 480
QY 481 GAAGCCCAACCAATTTGATCTTGAAACTTTGGAAATGATACACAAAACAAACAGGAGCA 540
Db 481 GAAGCCCAACCAATTTGATCTTGAAACTTTGGAAATGATACACAAAACAAACAGGAGCA 540

RESULT 5
AAS51602
ID AAS51602 standard; DNA; 864 BP.
XX AC AAS51602;
XX DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #19.
XX Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.
XX WO200170955-A2.
PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.

(ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR P-PSDB; AAU33743.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 4184; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 864 BP; 323 A; 112 C; 182 G; 247 T; 0 other;

Query Match 55.1%; Score 474; DB 23; Length 864;
Best Local Similarity 99.8%; Pred. No. 1.9e-225;
Matches 524; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 ATGAATAAATTAATAGATCAAGTCAATTAATGAATTTATCGGTTGCGATAAATAATCAAGTA 75
Db 1 ATGAATAAATTAATAGATCAAGTCAATTAATGAATTTATCGGTTGCGATAAATAATCAAGTA 60
QY 76 ATGGATCTACGCTAGAGAAAAGTATGTTGTTATTCATTAATGCTGGAGGTAAACGCATC 135
Db 61 ATGGATCTACGCTAGAGAAAAGTATGTTGTTATTCATTAATGCTGGAGGTAAACGCATC 120
QY 136 CGACCAAGTTCTGTTTATTTACTTCTACTTTAGATTACCAATAATACCGAGTATGAGTAGTATG 195
Db 121 CGACCAAGTTCTGTTTATTTACTTCTACTTTAGATTACCAATAATACCGAGTATGAGTAGTATG 180
QY 196 AAGACGCAATTTGCACCTAGATAATGATTCATACATATTCTACTTATTCATGATGACCTACCA 255
Db 181 AAGACGCAATTTGCACCTAGATAATGATTCATACATATTCTACTTATTCATGATGACCTACCA 240
QY 256 GCGATGGATAATGATGATGATTCGACGAGGAAAATTAACAATAATCAATAAGTATATGTTGAG 315
Db 241 GCGATGGATAATGATGATGATTCGACGAGGAAAATTAACAATAATCAATAAGTATATGTTGAG 300
QY 316 TGGACTGCGAATTTAGCAGGTGATGCTTTTATTAACATAAGCAATTTTCAAGT 375
Db 301 TGGACTGCGAATTTAGCAGGTGATGCTTTTATTAACATAAGCAATTTTCAAGT 360
QY 376 GATGATAGATTAACTGATCAAGTAAATAAAGTTCTACACGGCTGTCATACCAAGT 435
Db 361 GATGATAGATTAACTGATCAAGTAAATAAAGTTCTACACGGCTGTCATACCAAGT 420
QY 436 GGTCAATGTTGGAATGTCGCGGTCAAATGTTAGATATGCAAAAGCGAAGGCCAACCAATT 495
Db 421 GGTCAATGTTGGAATGTCGCGGTCAAATGTTAGATATGCAAAAGCGAAGGCCAACCAATT 480
QY 496 GATCTTGAACCTTTGGAATGATACACAAAACAAACAGGAGCA 540
Db 481 GATCTTGAACCTTTGGAATGATACACAAAACAAACAGGAGCA 525

RESULT 6
AAS50245/c
ID AAS50245 standard; DNA; 413 BP.
XX AC AAS50245;
XX DT 13-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation inhibitory sequence #1469.

XX Antisense; ss; prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

XX 27-SEP-2001.

XX	21-MAR-2001; 2001WO-US09180.	
PF		
XX	21-MAR-2000; 2000US-191078P.	
XX	23-MAY-2000; 2000US-206848P.	
PR	26-MAY-2000; 2000US-207727P.	
PR	26-MAY-2000; 2000US-207727P.	
PR	23-OCT-2000; 2000US-242578P.	
PR	27-NOV-2000; 2000US-253625P.	
PR	28-DEC-2000; 2000US-257931P.	
PR	16-FEB-2001; 2001US-269308P.	
XX	(ELIT-) ELITRA PHARM INC.	
PA		
XX	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;	
FI	Yamamoto RT, Xu HH;	
PI	WPI; 2001-611495/70.	
XX		
DR		
XX		
PT	New polynucleotides for the identification and development of	
	antibiotics, comprise sequences of antisense nucleic acids -	
XX	Claim 1; Seq ID No 2822; 51lpp; English.	
CC	The invention relates to antisense inhibitors of genes essential to	
CC	prokaryotic cellular proliferation, their use in identifying the	
CC	genes, themselves and the encoded proteins. The prokaryotes used are	
CC	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella	
CC	pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The	
CC	invention is also useful for the identification of potential new targets	
CC	for antibiotic development. The antisense nucleic acids can also be used	
CC	to identify proteins used in proliferation, to express these proteins,	
CC	and to obtain antibodies capable of binding to the expressed proteins.	
CC	The proteins can be used to screen compounds in rational drug discovery	
CC	programmes. The antisense nucleic acid sequence is also useful to screen	
CC	for homologous nucleic acids which are required for cell proliferation in	
CC	a wide variety of organisms. The present sequence is an antisense	
CC	oligonucleotide of the invention.	
CC	Note: The sequence data for this patent did not form part	
CC	of the printed specification, but was obtained in electronic	
CC	format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences.	
XX		
XX	Sequence 413 BP; 115 A; 89 C; 58 G; 151 T; 0 other;	
XX		
XX	Query Match 48.0%; Score 413; DB 23; Length 413;	
XX	Best Local Similarity 100.0%; Pred. No. 3.9e-195;	
XX	Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
XX		
Db	160 TTGATTTCACATAATACCGAGTAGTAGTATGAAGACGCCAATTGCAC TAGAATG 219	
Db	413 TTGATTTCACATAATACCGAGTAGTAGTATGAAGACGCCAATTGCAC TAGAATG 354	
Qy	220 ATTCATACATATTCACATTATTCATGATGACCTACCGAGTGGATAATGATGATTATTCGA 279	
Db	353 ATTCATACATATTCACATTATTCATGATGACCTACCGAGTGGATAATGATGATTATTCGA 294	
Qy	280 CGAGGAAAAATTAACAAATCATATAAGTATATGGTGGACTGGCATATTAGCAGGTGAT 339	
Db	293 CGAGGAAAAATTAACAAATCATATAAGTATATGGTGGACTGGCATATTAGCAGGTGAT 234	
Qy	340 GCATTATTAACTAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 399	
Db	233 GCATTATTAACTAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174	
Qy	400 AAAATAAAAGTCTTACAAAGCGCTGCTCAATAGCAAGTGGTCTATGTTGGAATGGTCGGCGGT 459	
Db	173 AAAATAAAAGTCTTACAAAGCGCTGCTCAATAGCAAGTGGTCTATGTTGGAATGGTCGGCGGT 114	
Qy	460 CAAATGTTAGATATGCAAAAGCGAAGCCAACTTATGATCTTGAACACTTTCGGAATGATA 519	
Db	113 CAAATGTTAGATATGCAAAAGCGAAGCCAACTTATGATCTTGAACACTTTCGGAATGATA 54	

Qy	520	CACAAACAAAACAGAGCATTATTAACTTTTGCAGTTATGAGTGCAGACA	572
Db	53	CACAAACAAAACAGAGCATTATTAACTTTTGCAGTTATGAGTGCAGACA	1
RESULT 7			
AAS50698/c			
ID	AAS50698 standard; DNA; 337 BP.		
XX	AC	AAS50698;	
XX	AC		
XX	DT	13-FEB-2002 (first entry)	
XX	XX	Staphylococcus aureus cellular proliferation inhibitory sequence #	
DE	XX	Antisense; ss; prokaryotic cellular proliferation;	
KW	KW	antibiotic; antibacterial; drug design.	
KW	XX		
OS	XX	Staphylococcus aureus.	
XX	XX	WO200170955-A2.	
PN	XX		
PD	XX	27-SEP-2001.	
XX	XX		
PF	XX	21-MAR-2001; 2001WO-US09180.	
XX	XX		
PR	XX	21-MAR-2000; 2000US-191078P.	
PR	XX	23-MAY-2000; 2000US-206848P.	
PR	XX	26-MAY-2000; 2000US-207737P.	
PR	XX	23-OCT-2000; 2000US-242578P.	
PR	XX	27-NOV-2000; 2000US-253625P.	
PR	XX	22-DEC-2000; 2000US-257931P.	
XX	XX	16-FEB-2001; 2001US-269308P.	
PA	XX	(ELIT-) ELITRA PHARM INC.	
XX	XX		
PI	XX	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr	
PI	XX	Yamamoto RT, Xu HH;	
XX	XX	WPI; 2001-611495/70.	
XX	XX		
PT	XX	New polynucleotides for the identification and development of	
PT	XX	antibiotics, comprise sequences of antisense nucleic acids -	
XX	XX		
PS	XX	Claim 1; Seq ID No 3275; 511pp; English.	
XX	XX		
CC	CC	The invention relates to antisense inhibitors of genes essential to	
CC	CC	prokaryotic cellular proliferation, their use in identifying the	
CC	CC	genes, their use in the discovery of novel antibiotics, the essen-	
CC	CC	genes themselves and the encoded proteins. The prokaryotes used are	
CC	CC	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsie-	
CC	CC	lla pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The	
CC	CC	invention is also useful for the identification of potential new tar-	
CC	CC	gets for antibiotic development. The antisense nucleic acids can also be	
CC	CC	used to identify proteins used in proliferation, to express these protei-	
CC	CC	ns and to obtain antibodies capable of binding to the expressed protei-	
CC	CC	ns. The proteins can be used to screen compounds in rational drug discov-	
CC	CC	ery programmes. The antisense nucleic acid sequence is also useful to	
CC	CC	for homologous nucleic acids which are required for cell prolifera-	
CC	CC	tion in a wide variety of organisms. The present sequence is an antisense	
CC	CC	oligonucleotide of the invention.	
CC	CC		
CC	CC	Note: The sequence data for this patent did not form part	
CC	CC	of the printed specification, but was obtained in electronic	
CC	CC	format directly from WIPO at	
CC	CC	ftp.wipo.int/pub/published_pct_sequences.	

QY	460	CAAAATGTTAGATATGCAAAAGCGAAGGCCAACCAATTGATCTTGAAACCTTTGGAAATGATA	519
Db	113	CAAAATGTTAGATATGCAAAAGCGAAGGCCAACCAATTGATCTTGAAACCTTTGGAAATGATA	54


```
Db 308 GAGGAAATTAACAAATCATAAGT 332
|||||
RESULT 10
AAH54780
ID AAH54780 standard; DNA; 3444 BP.
XX
AC AAH54780;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:4144.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
XX vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 8; Page 1844-1845; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis,
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 3444 BP; 1365 A; 451 C; 571 G; 1057 T; 0 other;
Query Match 2.9%; Score 25; DB 22; Length 3444;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GAGGAAATTAACAAATCATAAGT 305
Db 1046 GAGGAAATTAACAAATCATAAGT 1070
|||||
RESULT 11
AAH54634/C
ID AAH54634 standard; DNA; 3454 BP.
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3929.
XX
```

```
XX
AC AAH54634;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3998.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
XX vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 8; Page 1671-1672; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis,
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 3454 BP; 1005 A; 574 C; 493 G; 1382 T; 0 other;
Query Match 2.9%; Score 25; DB 22; Length 3454;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GAGGAAATTAACAAATCATAAGT 305
Db 77 GAGGAAATTAACAAATCATAAGT 53
|||||
RESULT 12
AAH54565/C
ID AAH54565 standard; DNA; 4045 BP.
XX
AC AAH54565;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3929.
XX
```


KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis; ds.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

XX 17-MAY-2001.

PD 09-NOV-2000; 2000WO-US30782.

XX 09-NOV-1999; 99US-0164258.

XX (GLAX) GLAXO GROUP LTD.

PA Kimmerly WJ;

PI WPI; 2001-316495/33.

DR Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX useful for vaccinating against infections, e.g. endocarditis -

PS Claim 8; Page 1586-1588; 2188pp; English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAH81454 to AAH83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.

CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 4045 BP; 1249 A; 691 C; 536 G; 1569 T; 0 other;

Query Match 2.9%; Score 25; DB 22; Length 4045;

Best Local Similarity 100.0%; Pred. No. 0.027;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GAGGAAATTAACAATCATTAAGT 305

Db 2033 GAGGAAATTAACAATCATTAAGT 2009

RESULT 13

ABA90521

ID ABA90521 standard; DNA; 2365589 BP.

XX ABA90521;

XX 16-MAY-2002 (first entry)

DE Genomic sequence of Lactococcus lactis IL1403.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.

OS Lactococcus lactis IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000FR-0004630.

XX 11-APR-2000; 2000FR-0004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species -

PS Claim 1; SEQ ID 1; 2504pp; French.

CC The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: the sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;

Query Match 2.3%; Score 20; DB 24; Length 2365589;

Best Local Similarity 100.0%; Pred. No. 7.7;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 ATGGATAATGATGATTATCG 278

Db 881007 ATGGATAATGATGATTATCG 881026

RESULT 14

ABA74941/c

ID ABA74941 standard; DNA; 151 BP.

XX ABA74941;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #23246.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX Claim 4; SEQ ID NO 23246; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single-exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 151 BP; 31 A; 22 C; 24 G; 74 T; 0 other;

Query Match 2.2%; Score 19; DB 22; Length 151;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

503 AACCTTTGGAATGATACA 521

|||||

139 AACCTTTGGAATGATACA 121

RESULT 15

ABA39630/c

ID ABA39630 standard; DNA; 151 BP.

XX

AC ABA39630;

XX

DT 23-JAN-2002 (first entry)

XX

DE Probe #18096 for gene expression analysis in human heart cell sample.

XX

KW Human; gene expression; heart; microarray; vascular system; probe;

KW

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW

XX congenital heart disease; ss.

OS

XX Homo sapiens.

XX

XX WO200157274-A2.

PN

PD 09-AUG-2001.

XX

XX 30-JAN-2001; 2001WO-US00666.

XX

XX 04-FEB-2000; 2000US-0180312.

XX

XX 26-MAY-2000; 2000US-0207456.

XX

XX 30-JUN-2000; 2000US-0608408.

PR

PR 03-AUG-2000; 2000US-0632366.

PR

PR 21-SEP-2000; 2000US-0234687.

PR

PR 27-SEP-2000; 2000US-0236359.

PR

XX 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX

XX WPI; 2001-488899/53.

DR

XX Single exon nucleic acid probes for analyzing gene expression in human

XX

XX hearts -

XX

PS Claim 4; SEQ ID No 18096; 530pp; English.

XX

XX The present invention relates to single exon nucleic acid probes for

XX

XX measuring human gene expression in a sample derived from human heart. The

XX

XX present sequence is one such probe. The probes may be used for

XX

XX predicting, measuring and displaying gene expression in samples derived

XX

XX from the human heart via microarrays. By measuring gene expression, the

XX

XX probes are useful for predicting, diagnosing, grading, staging,

CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
XX Sequence 151 BP; 31 A; 22 C; 24 G; 74 T; 0 other;

Query Match 2.2%; Score 19; DB 22; Length 151;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 AACCTTTGGAATGATACA 521

|||||

Db 139 AACCTTTGGAATGATACA 121

Search completed: May 30, 2003, 18:26:12

Job time : 283 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 17:36:25 ; Search time 61 Seconds
(without alignments)
4328.667 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atacgaatcaccgatgaa.....tattagaatcgttgattta 861

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	606	70.4	882	3	US-09-276-873-1
2	25	2.9	909	4	Sequence 1, Appli
3	18	2.1	763	4	Sequence 334, App
4	18	2.1	1301	1	Sequence 110, App
5	18	2.1	1301	1	Sequence 1, Appli
6	18	2.1	5163	3	Sequence 1, Appli
7	18	2.1	5163	3	Sequence 4, Appli
8	18	2.1	5318	2	Sequence 3, Appli
9	18	2.1	5318	3	Sequence 3, Appli
10	17	2.0	604	4	Sequence 404, App
11	17	2.0	616	4	Sequence 173, App
12	17	2.0	660	2	Sequence 1, Appli
13	17	2.0	1164	4	Sequence 2199, Ap
14	17	2.0	1511	2	Sequence 3, Appli
15	17	2.0	1511	3	Sequence 3, Appli
16	17	2.0	2335	2	Sequence 3, Appli
17	17	2.0	2335	3	Sequence 3, Appli
18	17	2.0	4656	4	Sequence 3, Appli
19	17	2.0	6614	4	Sequence 4, Appli
20	17	2.0	7358	4	Sequence 8, Appli
c 21	17	2.0	4403765	4	Sequence 2, Appli
c 22	17	2.0	4411529	4	Sequence 1, Appli
23	16	1.9	90	4	Sequence 7, Appli
c 24	16	1.9	384	4	Sequence 25, Appli
c 25	16	1.9	507	4	Sequence 2678, Ap
26	16	1.9	621	4	Sequence 1593, Ap
27	16	1.9	690	4	Sequence 2192, Ap

28	16	1.9	732	4	US-09-134-001C-1402	Sequence 1402, Ap
29	16	1.9	861	3	US-09-187-050-4	Sequence 4, Appli
30	16	1.9	863	1	US-08-592-126-71	Sequence 71, Appli
31	16	1.9	885	3	US-09-187-050-11	Sequence 11, Appli
32	16	1.9	909	2	US-08-286-819A-22	Sequence 22, Appli
33	16	1.9	909	3	US-08-980-357-22	Sequence 22, Appli
34	16	1.9	980	4	US-09-171-209-8	Sequence 8, Appli
35	16	1.9	1179	3	US-09-187-050-13	Sequence 13, Appli
36	16	1.9	1179	3	US-09-187-050-15	Sequence 15, Appli
37	16	1.9	1179	3	US-09-187-050-17	Sequence 17, Appli
38	16	1.9	1179	3	US-09-187-050-19	Sequence 19, Appli
39	16	1.9	1179	3	US-09-187-050-21	Sequence 21, Appli
40	16	1.9	1179	3	US-09-187-050-23	Sequence 23, Appli
41	16	1.9	1179	3	US-09-187-050-25	Sequence 25, Appli
42	16	1.9	1378	1	US-08-592-126-72	Sequence 72, Appli
c 43	16	1.9	1508	3	US-08-714-918-31	Sequence 31, Appli
c 44	16	1.9	1508	4	US-09-265-315-31	Sequence 31, Appli
c 45	16	1.9	1508	4	US-09-265-315-31	Sequence 31, Appli

ALIGNMENTS

RESULT 1
US-09-276-873-1
; Sequence 1, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: Ispa
; FILE REFERENCE: GMI0208
; CURRENT APPLICATION NUMBER: US/09/276.873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-276-873-1

Query Match 70.4%; Score 606; DB 3; Length 882;
Best Local Similarity 99.4%; Pred. No. 2.7e-293;
Matches 856; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	ATGACGATCTACCGATGAATTAATAGATGAAGTCAATAATGAATTAATCGGTGCG	60
DB	1	ATGACGATCTACCGATGAATTAATAGATGAAGTCAATAATGAATTAATCGGTGCG	60
QY	61	ATAAATAATCAGTAATGATGATCAGCTAGAGAAGTATGTTGTTTCAATTAATGCT	120
DB	61	ATAAATAATCAGTAATGATGATCAGCTAGAGAAGTATGTTGTTTCAATTAATGCT	120
QY	121	GGAGTTAAACGATCCGACCGAGTCTGTTTATCTACTCATTAGATTCATTAATACCGAG	180
DB	121	GGAGTTAAACGATCCGACCGAGTCTGTTTATCTACTCATTAGATTCATTAATACCGAG	180
QY	181	TATGATTTAGTGTATGAAGGCGCAATTCACCTAGAAAATGATTCATACATATTCATCTATT	240
DB	181	TATGATTTAGTGTATGAAGGCGCAATTCACCTAGAAAATGATTCATACATATTCATCTATT	240
QY	241	CATGATGACCTACCGACCGATGATTAATGATGATTCATGACGAGGAGAAATTAACAATCAT	300
DB	241	CATGATGACCTACCGACCGATGATTAATGATGATTCATGACGAGGAGAAATTAACAATCAT	300
QY	301	AAAGTATATGTTGAGTGCAGTCCGATATAGCAGGTGATGCTTTTATTAACTAACGATTT	360
DB	301	AAAGTATATGTTGAGTGCAGTCCGATATAGCAGGTGATGCTTTTATTAACTAACGATTT	360
QY	361	GAACCTTTTCAAGTGATGATGATTAATGATGATTAAGTAAATAAATTAAGTTCTCAACGG	420
DB	361	GAACCTTTTCAAGTGATGATGATTAATGATGATTAAGTAAATAAATTAAGTTCTCAACGG	420

QY 421 CTGTCATAGCAAGTGTCTGATTTGGAATGTCGGGGTCAAAATGTTAGATATGCAAGC 480
Db 421 CTGTCATAGCAAGTGTCTGATTTGGAATGTCGGGGTCAAAATGTTAGATATGCAAGC 480
QY 481 GAAGGCCAACCAATTTGATCTTTGAACTTTTGAATGATACACAAAACAGGAGCA 540
Db 481 GAAGGCCAACCAATTTGATCTTTGAACTTTTGAATGATACACAAAACAGGAGCA 540
QY 541 TTATTAACCTTTTGGGTTATGATGTCAGCAGATATCGCTAATGTCGATGATCAACTAAA 600
Db 541 TTATTAACCTTTTGGGTTATGATGTCAGCAGATATCGCTAATGTCGATGATCAACTAAA 600
QY 601 GAACATTTAGAAAGTTATAGTTATCATTTTATGATGTCAGATATGTCAGATTAAGATTTA 660
Db 601 GAACATTTAGAAAGTTATAGTTATCATTTTATGATGTCAGATATGTCAGATTAAGATTTA 660
QY 661 TTAGACTGCTATGCTGATGAAGCAAAAGTTAGGTAAAAAGTGGCGCAGCATCTTGAAT 720
Db 661 TTAGACTGCTATGCTGATGAAGCAAAAGTTAGGTAAAAAGTGGCGCAGCATCTTGAAT 720
QY 721 AATAAAGTACGTACGTGAGTTTATAGGAAAGATGGCGCAGATTAATTTGACTTAT 780
Db 721 AATAAAGTACGTACGTGAGTTTATAGGAAAGATGGCGCAGATTAATTTGACTTAT 780
QY 781 CATAGAGACGACGACGTGATGAACCAAAATTTGATCAACAAATTCATACAAAACAC 840
Db 781 CATAGAGACGACGACGTGATGAACCAAAATTTGATCAACAAATTCATACAAAACAC 840
QY 841 TTATTAGAAATCTTGTATTTA 861
Db 841 TTATTAGAAATCTTGTATTTA 861

RESULT 2

US-09-134-001C-334
; Sequence 334, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 334
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-334

Query Match 2.9%; Score 25; DB 4; Length 909;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 GAGGAAAATTAACAAATCAATCAAGT 305
Db 308 GAGGAAAATTAACAAATCAATCAAGT 332

RESULT 3

US-09-484-970B-110
; Sequence 110, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkumuth, Wayne
; APPLICANT: Walker, Michael G.

; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 110
; LENGTH: 763
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6426186 009051.5CB1
US-09-484-970B-110

Query Match 2.1%; Score 18; DB 4; Length 763;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GATGAATAAATTAATAGA 32
Db 212 GATGAATAAATTAATAGA 229

RESULT 4

US-07-686-322A-1
; Sequence 1, Application US/07686322A
; Patent No. 5312733
; GENERAL INFORMATION:
; APPLICANT: MacLeod Dr., Carol L.
; TITLE OF INVENTION: No. 5312733el T-cell Lymphoma cDNA Clones
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Department, Fulbright & Jaworski
; STREET: 1301 McKinney, Suit 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/686,322A
; FILING DATE: 19910411
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/509684
; FILING DATE: 13-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene A.
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: D-5232-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-3634
; TELEFAX: (713) 651-5246
; TELEX: Western Union 762829
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1301 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORGANISM: Mouse
; STRAIN: AKR1 Jackson
; INDIVIDUAL ISOLATE: SL12 cell line
; TISSUE TYPE: Lymphoma
; CELL TYPE: T-cell
; CELL LINE: SL12.3 and SL12.4

IMMEDIATE SOURCE:

CLONE: 19.1

US-07-686-322A-1

Query Match

Best Local Similarity 2.1%; Score 18; DB 1; Length 1301;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 828 CAATCAAAACACTTATT 845

Db 3 CAATCAAAACACTTATT 20

RESULT 5

US-08-002-999-1

Sequence 1, Application US/08002999

Patent No. 5440017

GENERAL INFORMATION:

APPLICANT: MacLeod Dr., Carol L.

TITLE OF INVENTION: No. 5440017el T-cell Lymphoma cDNA Clones

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

STREET: Patent Department, Fulbright & Jaworski

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77010-3095

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/002,999

FILING DATE: 19930111

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/686,322

FILING DATE: 11-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Adler, Benjamin A.

REGISTRATION NUMBER: 35,423

REFERENCE/DOCKET NUMBER: D-5232-DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (713) 651-5587

TELEFAX: (713) 651-5246

TELEX: Western Union 762829

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1301 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: Mouse

STRAIN: AKR1 Jackson

INDIVIDUAL ISOLATE: SL12 cell line

TISSUE TYPE: Lymphoma

CELL TYPE: T-cell

CELL LINE: SL12.3 and SL12.4

IMMEDIATE SOURCE:

CLONE: 19.1

US-08-002-999-1

Query Match

Best Local Similarity 2.1%; Score 18; DB 1; Length 1301;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 828 CAATCAAAACACTTATT 845

Db 3 CAATCAAAACACTTATT 20

RESULT 6

US-08-700-651-1

Sequence 1, Application US/08700651B

Patent No. 6015882

GENERAL INFORMATION:

APPLICANT: PETERSEN, CAROLYN

APPLICANT: LEECH, JAMES

APPLICANT: NELSON, RICHARD, C.

APPLICANT: GUT, JIRI

TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS

TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum

TITLE OF INVENTION: INFECTIONS

FILE REFERENCE: 480.19-4(HV)

CURRENT APPLICATION NUMBER: US/08/700,651B

CURRENT FILING DATE: 1997-08-14

EARLIER APPLICATION NUMBER: 08/415,751

EARLIER FILING DATE: 1995-04-03

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 5163

TYPE: DNA

ORGANISM: Cryptosporidium parvum

US-08-700-651-1

Query Match

Best Local Similarity 2.1%; Score 18; DB 3; Length 5163;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 490 CCAATTGATCTTGAAACT 507

Db 3769 CCAATTGATCTTGAAACT 3786

RESULT 7

US-08-928-361B-4

Sequence 4, Application US/08928361B

Patent No. 6071518

GENERAL INFORMATION:

APPLICANT: Petersen, Carolyn

TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

TITLE OF INVENTION: SPECIES INFECTIONS

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: PETERS, VERNY, JONES & BIKSA

STREET: 385 Sherman Avenue, Suite 6

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-1840

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,361B

FILING DATE: 12-SEP-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/026,062

FILING DATE: 13-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Veeriy, Hana

REGISTRATION NUMBER: 30,518

REFERENCE/DOCKET NUMBER: 480.76-1(HV)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-1677

TELEFAX: 650-324-1678

```
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5163 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-928-361B-4

Query Match          2.1%; Score 18; DB 3; Length 5163;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 490 CCAATTGATCTTGAAGT 507
Db 3769 CCAATTGATCTTGAAGT 3786

RESULT 8
US-08-928-651-2
Sequence 2, Application US/08700651B
Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-4 (HV)
; CURRENT APPLICATION NUMBER: US/08/700,651B
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 08/415,751
; EARLIER FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5318
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-08-700-651-2

Query Match          2.1%; Score 18; DB 3; Length 5318;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 490 CCAATTGATCTTGAAGT 507
Db 3769 CCAATTGATCTTGAAGT 3786

RESULT 9
US-08-928-361B-3
Sequence 3, Application US/08928361B
Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verny, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-928-361B-3

Query Match          2.1%; Score 18; DB 3; Length 5318;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 490 CCAATTGATCTTGAAGT 507
Db 3769 CCAATTGATCTTGAAGT 3786

RESULT 10
US-09-385-982-404
Sequence 404, Application US/09385982
Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(604)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-404

Query Match          2.0%; Score 17; DB 4; Length 604;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 CTCAGCTAGAGAAAGT 99
Db 289 CTCAGCTAGAGAAAGT 305

RESULT 11
```

```

1  EARLIER APPLICATION NUMBER: 60/117,393
2  EARLIER FILING DATE: 1999-01-27
3  EARLIER APPLICATION NUMBER: 60/098,639
4  EARLIER FILING DATE: 1998-08-31
5  NUMBER OF SEQ ID NOS: 544
6  SOFTWARE: FASTSEQ for Windows Version 3.0
7  SEQ ID NO 173
8  LENGTH: 616

```

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(616)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-173

```

```

? PRIOR FILLING DATE: 1997-08-14
? NUMBER OF SEQ ID NOS: 5674
? SEQ ID NO 2199
? LENGTH: 1164
? TYPE: DNA
? ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2199

```

US-09-134-001C-2199

RESULT 14
US-08-809-763-3
; Sequence 3, Application US/08809763
; Patent NO. 5919691

```

, STREET: 403 LEXINGTON AVENUE
, CITY: New York
, STATE: NY
, COUNTRY: USA
, ZIP: 10174
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Diskette
, COMPUTER: IBM Compatible
, OPERATING SYSTEM: DOS
, SOFTWARE: FastSeq for Windows
, CURRENT APPLICATION DATA:
, Version 2.0

```

APPLICATION NUMBER: US/08/809,763
FILING DATE: 21-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4257.204-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Acremonium sp.
STRAIN: CBS 265.95
08-809-763-3

Query Match 2.0%; Score 17; DB 2; Length 1511;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 489 ACCAATTGATCTTGAAA 505
|||||
DB 1436 ACCAATTGATCTTGAAA 1452

RESULT 15

US-08-956-253-3
Sequence 3, Application US/08956253
Patent No. 6071735
GENERAL INFORMATION:
APPLICANT: Andersen, Lene No. 6071735boe
APPLICANT: Lassen, S ren Flensted
APPLICANT: Kauppinen, Markus Sakari
APPLICANT: Nielsen, Jack Bech
TITLE OF INVENTION: An Enzyme and Enzyme Preparation
TITLE OF INVENTION: With Endoglucanase Activity
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6071735o No. 6071735disk of No. 6071735th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,253
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/809,763
FILING DATE: 21-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4257.204-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Acremonium sp.
STRAIN: CBS 265.95
US-08-956-253-3

Query Match 2.0%; Score 17; DB 3; Length 1511;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 489 ACCAATTGATCTTGAAA 505
|||||
DB 1436 ACCAATTGATCTTGAAA 1452

Search completed: May 30, 2003, 19:09:07
Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 19:07:56 ; Search time 186 Seconds
(without alignments)
6241.626 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atagcaatctacgatgaa.....tattagaatcgttgattta 861

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 845702 seqs, 674182571 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1691404

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCRTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	861	100.0	861	9	US-10-084-205-63
2	861	100.0	861	10	US-09-925-637-63
3	861	100.0	1893	7	US-08-781-986A-155
4	489	56.8	882	10	US-09-815-242-8485
5	474	55.1	864	10	US-09-815-242-4184
6	413	48.0	413	10	US-09-815-242-2822
7	337	39.1	337	10	US-09-815-242-3275
8	19	2.2	151	10	US-09-864-761-24950
9	19	2.2	600	10	US-09-864-761-8212
10	19	2.2	130427	9	US-10-175-523-87
11	19	2.2	640681	10	US-09-790-988-1
12	18	2.1	118	9	US-10-079-854-357
13	18	2.1	118	10	US-09-764-878-357
14	18	2.1	322	9	US-10-040-739-374
15	18	2.1	449	10	US-09-983-965-3145
16	18	2.1	721	9	US-09-764-868-109
17	18	2.1	1408	9	US-09-822-846-54
18	18	2.1	2064	9	US-09-822-846-55
19	18	2.1	6638	9	US-10-091-504-1645

c	20	18	2.1	6638	10	US-09-764-869-1645	Sequence 1645, Ap
c	21	18	2.1	10146	7	US-08-781-986A-243	Sequence 243, App
c	22	18	2.1	99014	10	US-09-880-107-3428	Sequence 3428, Ap
c	23	17	2.0	105	10	US-09-864-761-23287	Sequence 23287, A
c	24	17	2.0	189	10	US-09-783-590-5878	Sequence 5878, Ap
c	25	17	2.0	345	9	US-10-046-935-1794	Sequence 1794, Ap
c	26	17	2.0	345	9	US-09-878-178-1794	Sequence 1794, Ap
c	27	17	2.0	345	9	US-10-146-502-1794	Sequence 1794, Ap
c	28	17	2.0	394	10	US-09-880-107-139	Sequence 139, App
c	29	17	2.0	405	9	US-10-060-036-979	Sequence 979, App
c	30	17	2.0	472	10	US-09-864-761-6575	Sequence 6575, Ap
c	31	17	2.0	489	9	US-10-060-036-2589	Sequence 2589, Ap
c	32	17	2.0	498	9	US-09-918-995-23290	Sequence 23290, A
c	33	17	2.0	512	9	US-09-918-995-20190	Sequence 20190, A
c	34	17	2.0	523	9	US-10-060-036-1626	Sequence 1626, Ap
c	35	17	2.0	534	9	US-10-060-036-2901	Sequence 2901, Ap
c	36	17	2.0	540	9	US-09-736-457-1416	Sequence 1416, Ap
c	37	17	2.0	540	9	US-09-902-941-1416	Sequence 1416, Ap
c	38	17	2.0	540	9	US-09-849-626-1416	Sequence 1416, Ap
c	39	17	2.0	540	9	US-10-017-734-1416	Sequence 1416, Ap
c	40	17	2.0	566	10	US-09-864-761-13400	Sequence 13400, A
c	41	17	2.0	604	9	US-09-871-161-404	Sequence 404, App
c	42	17	2.0	616	9	US-09-871-161-173	Sequence 173, App
c	43	17	2.0	648	10	US-09-925-301-819	Sequence 819, App
c	44	17	2.0	660	10	US-09-847-185-1	Sequence 1, Appl
c	45	17	2.0	721	9	US-09-898-554-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-10-084-205-63
; Sequence 63, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-084-205-63

Query Match	100.0%	Score 861:	DB 9;	Length 861;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 861;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGACGAATCTACCGATGAATAAATAATAGATGAATGAATTAATCGTTTCGCG	60	
Db	1	ATGACGAATCTACCGATGAATAAATAATAGATGAATGAATTAATCGTTTCGCG	60	
Qy	61	ATAAATAAATCAAGTAATGGATCTACGCTAGAGAGAAAGTAGTTGTATTCAATTAATGCT	120	
Db	61	ATAAATAAATCAAGTAATGGATCTACGCTAGAGAGAAAGTAGTTGTATTCAATTAATGCT	120	
Qy	121	GGAGGTAACCGATCCGACCGAGTTCTGTATTACTCATTCTAGATTCACTAAATACCGAG	180	
Db	121	GGAGGTAACCGATCCGACCGAGTTCTGTATTACTCATTCTAGATTCACTAAATACCGAG	180	
Qy	181	TATCAGTTAGTATGAAGAGCGCAATTGCCTAGAGAAATGATTCATACATATTCTACTATT	240	
Db	181	TATCAGTTAGTATGAAGAGCGCAATTGCCTAGAGAAATGATTCATACATATTCTACTATT	240	

```
QY 241 CATGATGACCTACCGGATGGATTAATGATGATTCGACGAGGAAAAATTAACAATCAT 300
Db 241 CATGATGACCTACCGGATGGATTAATGATGATTCGACGAGGAAAAATTAACAATCAT 300
QY 301 AAAGTATATGCTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAGCATTT 360
Db 301 AAAGTATATGCTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAGCATTT 360
QY 361 GAACCTTATTTCAAGTATGATGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 420
Db 361 GAACCTTATTTCAAGTATGATGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 420
QY 421 CTGCTAATAGCAAGTGGTCAATGTTGAATGCTGCGCGGTCAAAATGTTAGTATGCAAGC 480
Db 421 CTGCTAATAGCAAGTGGTCAATGTTGAATGCTGCGCGGTCAAAATGTTAGTATGCAAGC 480
QY 481 GAAGGCCAACCAATTCATCTTGAACCTTTGGAATGATACACAAAACAAAACAGGACCA 540
Db 481 GAAGGCCAACCAATTCATCTTGAACCTTTGGAATGATACACAAAACAAAACAGGACCA 540
QY 541 TTATTAACTTTTTCGCTTATGATGCGAGGATATCGCTAATGTCGATGATACAACTAAA 600
Db 541 TTATTAACTTTTTCGCTTATGATGCGAGGATATCGCTAATGTCGATGATACAACTAAA 600
QY 601 GAACATTTAGAAAGTATAGTATTAATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 GAACATTTAGAAAGTATAGTATTAATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TTAGATGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 661 TTAGATGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 AATAAAGTACGTACGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 AATAAAGTACGTACGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 CATAGAGCGCAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 CATAGAGCGCAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TTATTAGAAATCGTTGATTTA 861
Db 841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2

US-09-925-637-63

Sequence 63, Application US/09925637

Patent No. US2002010338A1

GENERAL INFORMATION:

```
APPLICANT: Chohi
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
FILE REFERENCE: PB560
CURRENT APPLICATION NUMBER: US/09/925,637
CURRENT FILING DATE: 2001-08-10
PRIORITY APPLICATION NUMBER: PCT/US00/23773
PRIORITY FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US 60/151,933
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: US 08/781,986
PRIOR FILING DATE: 1997-01-03
PRIOR APPLICATION NUMBER: US 08/956,171
PRIOR FILING DATE: 1997-10-20
PRIOR APPLICATION NUMBER: US 60/009,861
PRIOR FILING DATE: 1996-01-06
NUMBER OF SEQ ID NOS: 74
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 63
LENGTH: 861
TYPE: DNA
ORGANISM: Homo sapiens
```

US-09-925-637-63

```
Query Match 100.0%; Score 861; DB 10; Length 861;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGAATCTACCGATGAATTAATTAAGTGAAGTCAATTAATGAATTAATCGGTGCG 60
Db 1 ATGACGAATCTACCGATGAATTAATTAAGTGAAGTCAATTAATGAATTAATCGGTGCG 60
QY 61 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
Db 61 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
QY 121 GGAGTAAACCGCATCCGACCGATTCCTGTTATTACTCAGTAAAGTAAAGTAAAGTAAAGT 180
Db 121 GGAGTAAACCGCATCCGACCGATTCCTGTTATTACTCAGTAAAGTAAAGTAAAGTAAAGT 180
QY 181 TATGAGTTAGGTATGAGAGCGCAATTCGACTAGAAATGATTCATATCATATTCACATTAT 240
Db 181 TATGAGTTAGGTATGAGAGCGCAATTCGACTAGAAATGATTCATATCATATTCACATTAT 240
QY 241 CATGATGACCTACCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 241 CATGATGACCTACCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 AAAGTATATGCTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAGCATTT 360
Db 301 AAAGTATATGCTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAGCATTT 360
QY 361 GAACCTTATTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 GAACCTTATTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 CTGCTAATAGCAAGTGGTCAATGTTGAATGCTGCGCGGTCAAAATGTTAGTATGCAAGC 480
Db 421 CTGCTAATAGCAAGTGGTCAATGTTGAATGCTGCGCGGTCAAAATGTTAGTATGCAAGC 480
QY 481 GAAGGCCAACCAATTCATCTTGAACCTTTGGAATGATACACAAAACAAAACAGGACCA 540
Db 481 GAAGGCCAACCAATTCATCTTGAACCTTTGGAATGATACACAAAACAAAACAGGACCA 540
QY 541 TTATTAACTTTTTCGCTTATGATGCGAGGATATCGCTAATGTCGATGATACAACTAAA 600
Db 541 TTATTAACTTTTTCGCTTATGATGCGAGGATATCGCTAATGTCGATGATACAACTAAA 600
QY 601 GAACATTTAGAAAGTATAGTATTAATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 GAACATTTAGAAAGTATAGTATTAATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TTAGATGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 661 TTAGATGCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 AATAAAGTACGTACGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 AATAAAGTACGTACGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 CATAGAGCGCAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 CATAGAGCGCAGCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TTATTAGAAATCGTTGATTTA 861
Db 841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 3

US-08-781-986A-155/c

Sequence 155, Application US/08781986A

Publication No. US20030054436A1

GENERAL INFORMATION:

```
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Human Genome Sciences, Inc.
;; STREET: 9410 Key West Avenue
;; CITY: Rockville
;; STATE: Maryland
;; COUNTRY: USA
;; ZIP: 20850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;; COMPUTER: HP Vectra 486/33
;; OPERATING SYSTEM: MSDOS version 6.2
;; SOFTWARE: ASCII Text
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/781.986A
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Benson, Bob
;; REGISTRATION NUMBER: 30,446
;; REFERENCE/DOCKET NUMBER: PB248PP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 155:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1893 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
US-08-781-986A-155

Query Match
Best Local Similarity 100.0%; Score 861; DB 7; Length 1893;
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGAATCTACCGATGAATTAATATAGATGAAGTCAATATGAATTAATCGGTTCG 60
DB 1314 ATGACGAATCTACCGATGAATTAATATAGATGAAGTCAATATGAATTAATCGGTTCG 1255

QY 61 ATAAATAATCAGTAATGGATGACTCAGCTAGAGAAATATGTTGTTATTCATTAATGCT 120
DB 1254 ATAAATAATCAGTAATGGATGACTCAGCTAGAGAAATATGTTGTTATTCATTAATGCT 1195

QY 121 GGAGTAAACCGATCCGACCACTGCTGTTATCTACTCTAGATTCACCTAAATACCGAG 180
DB 1194 GGAGTAAACCGATCCGACCACTGCTGTTATCTACTCTAGATTCACCTAAATACCGAG 1135

DB 181 TATGAGTTAGTATGAAGCGCAATTCAGTATGATTCATACATATTCATCTATT 240
DB 1134 TATGAGTTAGTATGAAGCGCAATTCAGTATGATTCATACATATTCATCTATT 1075

QY 241 CATGATGACCTACCGATGATTAATGATTAATTCAGCGAGAAATTAACAATCAT 300
DB 1074 CATGATGACCTACCGATGATTAATGATTAATTCAGCGAGAAATTAACAATCAT 1015

QY 301 AAAGTATATGTTGAGTGGACTGCGATATAGCAGGTGATGCTTTATTAAGTAAAGCATTT 360
DB 1014 AAAGTATATGTTGAGTGGACTGCGATATAGCAGGTGATGCTTTATTAAGTAAAGCATTT 955

QY 361 GAACCTATTTCAGGTGATGATGATTAATCTGATGAAGTAAATAAATTAAGTCTCAACGG 420
DB 954 GAACCTATTTCAGGTGATGATGATTAATCTGATGAAGTAAATAAATTAAGTCTCAACGG 895

QY 421 CTGTCATATAGCAAGTGGTCATGTTGGAATGTTGGAATGTTGCGCGGTCAAAATGATATGCAAGC 480
DB 894 CTGTCATATAGCAAGTGGTCATGTTGGAATGTTGGAATGTTGCGCGGTCAAAATGATATGCAAGC 835

QY 481 GAAGGCCAACCAATTTGATCTTTGAAACTTTGGAATGATACACAAAAACAAACAGGAGCA 540
DB 834 GAAGGCCAACCAATTTGATCTTTGAAACTTTGGAATGATACACAAAAACAAACAGGAGCA 775

RESULT 4

US-09-815-242-8485
; Sequence 8485, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 8485
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(882)
US-09-815-242-8485

Query Match
Best Local Similarity 56.8%; Score 489; DB 10; Length 882;
Matches 539; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 24950
LENGTH: 151
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL117667.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
OTHER INFORMATION: EST HUMAN HIT: AA601023.1, EVALUATE 1.00e-14
OTHER INFORMATION: NT HIT: AJ271735.1, EVALUATE 2.00e-15
OTHER INFORMATION: SWISSPROT HIT: P08348, EVALUATE 1.00e-05
US-09-864-761-24950

Query Match 2.2%; Score 19; DB 10; Length 151;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 AACCTTGGAAATGATACA 521
|||||
Db 139 AACCTTGGAAATGATACA 121

ULT 9
09-864-761-8212/c
Sequence 8212, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 8212
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL117667.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
US-09-864-761-8212

Query Match 2.2%; Score 19; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 AACCTTGGAAATGATACA 521
|||||
Db 470 AACCTTGGAAATGATACA 452

RESULT 10

US-10-175-523-87
Sequence 87, Application US/10175523
Publication No. US20030096264A1
GENERAL INFORMATION:
APPLICANT: Brockman, Jeffrey
APPLICANT: Evans, David
APPLICANT: Hook, Derek
APPLICANT: Klimczak, Leszek
APPLICANT: Laeng, Pascal
APPLICANT: Palfreyman, Michael
APPLICANT: Rajan, Prithi
TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
FILE REFERENCE: 3235/17795-US3
CURRENT APPLICATION NUMBER: US/10/175,523
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: US 60/299,151
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: US 60/317,828
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/325,150

; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 87
; LENGTH: 130427
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-079-854-357

Query Match 2.2%; Score 19; DB 9; Length 130427;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 647 TTAAGATGATTATTAGA 665
|||||
106008 TTAAGATGATTATTAGA 106026

RESULT 11
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENORU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 2.2%; Score 19; DB 10; Length 640681;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 AATRAATTAATAGATGAAG 37
|||||
DB 214082 AATRAATTAATAGATGAAG 214064

RESULT 12
US-10-079-854-357
; Sequence 357, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21C1
; CURRENT APPLICATION NUMBER: US/10/079,854
; CURRENT FILING DATE: 2002-02-22
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 357
; LENGTH: 118
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-079-854-357

Query Match 2.1%; Score 18; DB 9; Length 118;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 601 GAACATTTAGAAAGTTAT 618
|||||
DB 14 GAACATTTAGAAAGTTAT 31

RESULT 13
US-09-764-878-357
; Sequence 357, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 357
; LENGTH: 118
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-878-357

Query Match 2.1%; Score 18; DB 10; Length 118;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 601 GAACATTTAGAAAGTTAT 618
|||||
DB 14 GAACATTTAGAAAGTTAT 31

RESULT 14
US-10-040-739-374/c
; Sequence 374, Application US/10040739
; Patent No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John
; Lavallie, Edward
; Racie, Lisa
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/040,739
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:

; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 374:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 322 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 374:
US-10-040-739-374

Query Match 2.1%; Score 18; DB 9; Length 322;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
645 GATTAAAGATGATTTATT 662
|||||
271 GATTAAAGATGATTTATT 254

RESULT 15
US-09-983-965-3145/c
; Sequence 3145, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 3145
; TYPE: DNA
; LENGTH: 449
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 26-LIB3058-004-Q1-K1-G5
09-983-965-3145

Query Match 2.1%; Score 18; DB 10; Length 449;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 291 AACAAATCATAAAGTATA 308
|||||
Db 164 AACAAATCATAAAGTATA 147

Search completed: May 30, 2003, 20:32:14
Job time : 188 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 18:18:06 ; Search time 1599 Seconds
(without alignments)
8720.647 Million cell updates/sec

Title: US-09-925-637-63

Perfect score: 861

Sequence: 1 atgacgaatctaccgatgaa.....tattagaatcgttgattta 861

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	2.8	317	10	BB554288
C 2	22	2.6	415	17	AZ481013
C 3	22	2.6	631	10	AV679322
C 4	22	2.6	652	17	AZ506985
C 5	21	2.4	515	12	BG817859
C 6	21	2.4	1056	17	CNS060DRD

C 7	20	2.3	24	17	AZ852748
C 8	20	2.3	159	17	AQ355799
C 9	20	2.3	216	10	BB323434
10	20	2.3	250	10	BB323699
11	20	2.3	301	10	BB188250
C 12	20	2.3	427	10	AW986250
C 13	20	2.3	433	17	AQ380942
C 14	20	2.3	440	17	BH765020
15	20	2.3	444	10	BB822482
16	20	2.3	449	10	BB828059
C 17	20	2.3	456	14	C87470
C 18	20	2.3	475	10	BE210383
C 19	20	2.3	522	17	AQ256183
C 20	20	2.3	527	12	BG067198
21	20	2.3	547	10	BB821485
22	20	2.3	571	17	AZ989932
C 23	20	2.3	579	14	C88106
C 24	20	2.3	581	14	C85477
C 25	20	2.3	653	17	AQ352899
C 26	20	2.3	695	17	BH664215
C 27	20	2.3	704	17	BH546103
C 28	20	2.3	734	14	BQ031102
C 29	20	2.3	748	17	AZ175417
C 30	20	2.3	840	17	BH573787
C 31	20	2.3	850	17	AZ528307
32	20	2.3	1079	17	CNS05E6G
33	20	2.3	1362	12	BG335622
C 34	19	2.2	114	14	BQ979812
C 35	19	2.2	175	17	AL765391
36	19	2.2	187	10	AW890057
C 37	19	2.2	203	14	BQ977827
C 38	19	2.2	217	14	BQ978811
C 39	19	2.2	217	14	BQ979209
C 40	19	2.2	222	14	BQ973041
C 41	19	2.2	229	14	BT034345
C 42	19	2.2	230	14	BQ979463
C 43	19	2.2	232	14	BQ972776
C 44	19	2.2	233	10	AW430406
C 45	19	2.2	244	9	AV246789

ALIGNMENTS

RESULT 1
BB554288
LOCUS
DEFINITION
BB554288 RIKEN full-length enriched, 2 days pregnant adult female
ovary Mus musculus cDNA clone E330012K12 3' similar to J04596 Mouse
platelet-derived growth factor-inducible KC protein mRNA, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BB554288 317 bp mRNA linear EST 01-AUG-2000
BB554288 RIKEN full-length enriched, 2 days pregnant adult female
ovary Mus musculus cDNA clone E330012K12 3' similar to J04596 Mouse
platelet-derived growth factor-inducible KC protein mRNA, mRNA
sequence.
BB554288.1 GI:9640654
EST.
house mouse.
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Konno H., Akahira, S., Akiyama, J., Arakawa, T., Carninci
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Iizawa, M., Kadota, K., Kagawa, J., Ishikawa, J., Ishikawa, T., Itoh, M.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Shibata, K., Shibata
Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suganara, Y.,
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya
T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamana, I.,
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino
M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)

REFERENCE
1 (bases 1 to 631)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in *Ciona intestinalis*
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department Of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..631
Location/Qualifiers
/organism="Ciona intestinalis"
/db_xref="taxon:7719"
/clone="rcitb15k5"
/clone_lib="Nori Satoh unpublished cDNA library"
/tissue_type="whole animal"
/dev_stage="tailbud"
/note="Vector: pBluescript SK"
BASE COUNT 204 a 119 c 125 g 183 t
ORIGIN

Query Match 2.6%; Score 22; DB 10; Length 631;
Best Local Similarity 100.0%; Pred. No. 7.2; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 598 AAAGAACATTTAGAAAGTTATA 619
|||||
DB 171 AAAGAACATTTAGAAAGTTATA 150

RESULT 4
AZ506985 652 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0348H23F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION clone UUCG1M0348H23 F, DNA sequence.
ACCESSION AZ506985
VERSION AZ506985.1 GI:10688301
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 652)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0348 row: H column: 23
Seq primer: CGTTGTAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 652.

FEATURES
source
1..652
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6j"
/db_xref="taxon:10090"
/clone="UUCG1M0348H23"
/clone_lib="Mouse 10kb plasmid UUCG1M library"

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6j (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (GI:4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 204 a 143 c 128 g 177 t
ORIGIN

Query Match 2.6%; Score 22; DB 17; Length 652;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 AAAGATGATTATTAGACTGCT 670
|||||
DB 54 AAAGATGATTATTAGACTGCT 75

RESULT 5
BG817859/c 515 bp mRNA linear EST 22-MAY-2001
LOCUS ESS00281 S.scabiei cDNA library Sarcopes scabiei cDNA clone
DEFINITION SAS0368 5', mRNA sequence.
ACCESSION BG817859
VERSION BG817859.1 GI:14188839
KEYWORDS EST.
SOURCE Sarcopes scabiei.
ORGANISM Sarcopes scabiei
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Acariiformes; Sarcopiformes; Astigmata; Sarcoptoidea; Sarcoptidae;
Sarcopes.

REFERENCE 1 (bases 1 to 515)
AUTHORS Ljunggren,E.L., Nilsson,D., Naslund,K. and Mattsson,J.G.
TITLE Expressed sequence tag analysis of the parasitic mite Sarcopes
scabiei
JOURNAL Unpublished (2001)
COMMENT Contact: Mattsson J.G.
Department of Parasitology (SWEPAR)
National Veterinary Institute
SE-751 89 Uppsala, Sweden
Tel: +46 18 674120
Fax: +46 18
Email: jens.mattsson@sva.se
Seq primer: T3 primer
High quality sequence stop: 515.

FEATURES
source
1..515
Location/Qualifiers
/organism="Sarcopes scabiei"
/db_xref="taxon:52283"
/clone="SAS0368"
/clone_lib="S.scabiei cDNA library"

/note="The Sarcopes scabiei mixed life-stage library was
constructed by Jens G Mattsson. cDNAs were synthesized
from poly(A)+ RNA by oligo d(T) priming, size-selected and
directionally cloned into the Uni-ZAP lambda vector
(Stratagene). The primary library was amplified on
XL1-Blue MRF' cells."
BASE COUNT 199 a 72 c 92 g 152 t

ORIGIN

Query Match 2.4%; Score 21; DB 12; Length 515;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 814 ATTGATGACAAATTCATATACA 934
|||||
Db 483 ATTGATGACAAATTCATATACA 463

RESULT 6
CNS06DRD/c

LOCUS
DEFINITION T7 end of clone AR0AA015E03 of library AR0AA from strain CBS 732 of Zygosaccharomyces rouxii, genomic survey sequence.

ACCESSION AL394175.1 GI:12144998

VERSION
KEYWORDS GSS; clone AR0AA015E03; RST AR0AA015E03CPI.

SOURCE Zygosaccharomyces rouxii.

ORGANISM Zygosaccharomyces rouxii.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Zygosaccharomycetes.

REFERENCE

AUTHORS Souciet,J.L., Algie,M., Artiguenave,F., Blandin,G.,

de Montigny,J., Dujon,B., Bon,E., Brottier,P., Casaregola,S.,
de Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toifano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.

TITLE Genomic exploration of the Hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE 20584711

PUBMED 11152876

REFERENCE

AUTHORS de Montigny,J., Straub,M., Potier,S., Tekala,F., Dujon,B.,

Wincker,P., Artiguenave,F. and Souciet,J.

TITLE Genomic exploration of the hemiascomycetous yeasts: 8.

JOURNAL Zygosaccharomyces rouxii

MEDLINE FEBS Lett. 487 (1), 52-55 (2000)

PUBMED 11152883

REFERENCE

AUTHORS Genoscope.

JOURNAL Direct Submission

Submitted (28-MAR-2001) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

TITLE This BAC end sequence is part of a random genomic sequencing
program of thirteen yeast species:

. Saccharomyces bayanus var. uvarum,

. Saccharomyces exiguus,

. Saccharomyces servazzii,

. Zygosaccharomyces rouxii,

. Saccharomyces kluyveri,

. Kluyveromyces thermotolerans,

. Kluyveromyces lactis var. lactis,

. Kluyveromyces marxianus var. marxianus

. Pichia angusta,

. Debaryomyces hansenii var. hansenii,

. Pichia sorbitophila,

. Candida tropicalis and

. Yarrowia lipolytica.

Genomic inserts of 3 to 5 kb were prepared and both extremities
were sequenced. See keywords for description of this sequence and
for the sequence of the other extremity of this insert.

FEATURES

source

1. .1056

/organism="Zygosaccharomyces rouxii"

/strain="CBS 732"

/db_xref="taxon:4956"

/clone="AR0AA015E03"

/clone_lib="AR0AA"

/note="end : T7"

BASE COUNT 298 a 202 c 209 g 340 t 7 others

ORIGIN

Query Match 2.4%; Score 21; DB 17; Length 1056;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 GATGAACAAATTCATACAAA 837
|||||

Db 433 GATGAACAAATTCATACAAA 413

RESULT 7

AZ852748/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0155 row: G column: 09

Seq primer: CACACAGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 24.

Location/Qualifiers

1. .24

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0155G09"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 (gi14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into


```
/clone_lib="RIKEN full-length enriched, 4 days neonate
```

N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermotransformation and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki
Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

FEATURES

Location/Qualifiers

1. 301
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A330047P06"
/clone_lib="RIKEN full-length enriched, adult male spinal
cord"
/sex="male"
/tissue_type="spinal cord"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGATCCAGAGCGCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot - 10.0 and subtraction to Rot - 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATCTCGAGTAATTAATTAATCCCGCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

BASE COUNT 73 a 55 c 68 g 105 t
ORIGIN

Query Match 2.3%; Score 20; DB 10; Length 301;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

392 ATGAAGTAAATAAAGTT 411
|||||
277 ATGAAGTAAATAAAGTT 296

RESULT 12
AW986250/c
LOCUS AW986250 1427 bp mRNA linear EST 02-JUN-2000
DEFINITION IMAGE:1528006 3', mRNA sequence.

ACCESSION AW986250.1 GI:8180846
VERSION AW986250
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 427)

AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: uf75g12.y1
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LUNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:944106

High quality sequence stop: 363.

FEATURES

Location/Qualifiers

1. 427
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1528006"
/clone_lib="Soares mammary_gland_NLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 139 a 65 c 58 g 165 t
ORIGIN

Query Match 2.3%; Score 20; DB 10; Length 427;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 GTATGTTCTATTCATTAAT 117
|||||

Db 60 GTATGTTCTATTCATTAAT 41
|||||

RESULT 13
AQ380942/c
LOCUS AQ380942 433 bp DNA linear GSS 21-MAY-1999
DEFINITION RPII11-166B24-TV RPCI-11 Homo sapiens genomic clone RPCI-11-166B24,
DNA sequence.

ACCESSION AQ380942
VERSION AQ380942.1 GI:4351965
KEYWORDS GSS.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 433)

REFERENCE Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
AUTHORS J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building

JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: T7
Class: BAC ends.

Location/Qualifiers

1. 433
/organism="Homo sapiens"
/db_xref="GDB:7563407"
/db_xref="taxon:9606"
/clone="RPCI-11-166B24"

```

/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACE3.6; Site_1: ECORI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT      113 a   57 c   76 g   187 t
ORIGIN

Query Match      2.3%; Score 20; DB 17; Length 433;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 GAACTTTGGAATGATACA 521
Db 214 GAACTTTGGAATGATACA 195
|||||

RESULT 14
BH765020
US
INITIATION
BMBAC353B09T7_P5U Brugia malayi Genomic Bac Library 3 Brugia malayi
genomic, DNA sequence.
ACCESSION      BH765020
VERSION
KEYWORDS      GSS.
SOURCE        Brugia malayi.
ORGANISM      Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Brugia.
REFERENCE      1 (bases 1 to 440)
AUTHORS      Whittom,C., Daub,J., Ware,J., Quail,M., Hall,N., Barrell,B., Foster
,J., Guillano,D., Slatko,B. and Blaxter,M.
TITLE        Genome survey sequences from the human parasitic nematode Brugia
malayi
JOURNAL      Unpublished (2000)
COMMENT      Contact: Blaxter M.
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
Sequenced from the Brugia malayi BAC library constructed by Claire
Whittom and Dr Mike Quail. The sequence was generated by The
Pathogen Sequencing Unit, The Sanger Institute, Cambridge, UK in
collaboration with Mark Blaxter, ICAPB, University of Edinburgh,
Edinburgh, UK.
Seq primer: T7 (TAATAGGACTCACTATAGG)
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..440
/organism="Brugia malayi"
/strain="TRS"
/db_xref="taxon:6279"
/clone_lib="Brugia malayi Genomic Bac Library 3"
/sex="Mixed (male and female)"
/tissue_type="whole parasite"
/dev_stage="microfilaria (L1)"
/notes="Vector: pBACE3.6; Site_1: BamH I; Brugia malayi
genomic DNA was partially cleaved with Sau3A I and size
fractionated. 7,392 clones were generated with mean insert
size -48 kbp. The library was constructed by Claire
Whittom, Blaxter Nematode Genetics Lab, University of
Edinburgh, UK, and Dr Mike Quail, The Pathogen Sequencing
Unit, The Sanger Centre, Cambridge, UK."
BASE COUNT      151 a   55 c   56 g   178 t
ORIGIN

Query Match      2.3%; Score 20; DB 17; Length 440;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 359 TTGAACCTTTTCAACTGAT 378
Db 174 TTGAACCTTTTCAACTGAT 193
|||||

RESULT 15
BH822482
LOCUS
DEFINITION      BB822482
Jyg-MC(A) CDNA Mus musculus CDNA clone G830020C06 3', mRNA
sequence.
ACCESSION      BB822482
VERSION
KEYWORDS      EST.
SOURCE        house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 444)
AUTHORS      Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
,Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE        RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
JOURNAL      Unpublished (2001)
COMMENT      Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
FEATURES
source
Location/Qualifiers
1..444
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="G830020C06"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0526 Jyg-MC(A) CDNA"
/tissue_type="mammary gland"
/notes="pooled cell lines ; (cell_line=RCB-0526 Jyg-MC(A)"
(cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3),
(cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHPA),
(cell_line=RCB-0559 K-1 Fl), (cell_line=CRL-1283 B16
melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231

```


, (cell_type=leydig cells, cell_line=CRL-2065 MUTC-1),
 (tissue_type=bladder, cell_line=RCB-0544 MBr-2),
 (tissue_type=bone marrow, cell_type=stroma cell,
 cell_line= CRL-2028 SR-4987), (tissue_type=colon,
 cell_line=RCB-0549 Cle-H3), (tissue_type=kidney,
 cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
 cell_line=CRL-1734 SCA-9 clone 15), (strain=BA1B/C,
 cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3
), (strain=C3H, tissue_type=brain, cell_line=CRL-1443
 BC3H1)*

BASE COUNT 117 a 65 c 106 g 156 t
 ORIGIN

Query Match 2.3%; Score 20; DB 10; Length 444;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 392 ATGAAGTAAATAAAGTT 411

|||||

421 ATGAAGTAAATAAAGTT 440

Search completed: May 30, 2003, 19:36:01
 Job time : 1604 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 16:56:55 ; Search time 185 Seconds
(without alignments)
6275.364 Million cell updates/sec

Title: US-09-925-637-63

Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaatctgttgattta 861

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 845702 seqs, 674182571 residues

Total number of hits satisfying chosen parameters: 1691404

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	100.0	861	9	US-10-084-205-63
2	861	100.0	861	10	US-09-925-637-63
c 3	861	100.0	1893	7	US-08-781-986A-155
4	846.6	98.3	882	10	US-09-815-242-8485
5	831.6	96.6	864	10	US-09-815-242-4184
c 6	413	48.0	413	10	US-09-815-242-2822
c 7	337	39.1	337	10	US-09-815-242-3275
8	196.8	22.9	882	10	US-09-815-242-6533
9	195	22.6	7528	10	US-09-070-927A-55
10	164	19.0	811	10	US-09-974-300-1015
11	131.4	15.3	888	10	US-09-815-242-7142
12	124.8	14.5	876	10	US-09-815-242-9175
13	124.2	14.4	876	10	US-09-815-242-9499
c 14	114.2	13.3	640681	10	US-09-790-988-1
15	102.4	11.9	900	10	US-09-815-242-5972
16	98.6	11.5	891	9	US-09-941-947A-19
17	98.6	11.5	891	10	US-09-934-903-13
18	98.6	11.5	891	10	US-09-934-868-71
19	97.8	11.4	900	10	US-09-815-242-9986

20	89.6	10.4	936	10	US-09-815-242-4260	Sequence 4260, Ap
21	89.6	10.4	969	10	US-09-815-242-8505	Sequence 8505, Ap
22	88.6	10.3	819	9	US-10-084-205-65	Sequence 65, Appl
23	88.6	10.3	819	7	US-09-925-637-65	Sequence 65, Appl
24	88.6	10.3	884	7	US-08-781-986A-185	Sequence 185, App
c 25	79.6	9.2	498	10	US-09-070-927A-509	Sequence 509, App
26	77.4	9.0	498	10	US-09-974-300-5453	Sequence 5453, Ap
27	77.4	8.8	498	10	US-09-974-300-5438	Sequence 5438, Ap
c 28	75.2	8.7	1268	9	US-10-108-915-17	Sequence 17, Appl
c 29	73.4	8.5	2000	10	US-09-887-576-115	Sequence 115, App
30	63.8	7.4	1470	9	US-10-108-915-25	Sequence 25, Appl
31	60.8	7.1	912	10	US-09-815-242-7285	Sequence 7285, Ap
32	60.4	7.0	1062	9	US-10-108-915-15	Sequence 15, Appl
33	60	7.0	1161	9	US-10-108-915-37	Sequence 37, Appl
34	57.4	6.7	1441	9	US-10-108-915-21	Sequence 21, Appl
35	56.6	6.6	1131	10	US-09-934-778-1	Sequence 1, Appl
36	55.6	6.5	912	9	US-09-941-947A-25	Sequence 25, Appl
37	50.8	5.9	993	9	US-10-108-915-23	Sequence 23, Appl
38	49.6	5.8	912	10	US-09-815-242-7450	Sequence 7450, Ap
39	48.4	5.6	888	10	US-09-815-242-7873	Sequence 7873, Ap
40	47.8	5.6	623	9	US-10-108-915-19	Sequence 19, Appl
41	47.6	5.5	927	10	US-09-974-300-999	Sequence 999, App
42	47.6	5.5	972	10	US-09-815-242-6224	Sequence 6224, Ap
43	47.6	5.5	973	9	US-10-108-915-13	Sequence 13, Appl
44	44.6	5.2	474	10	US-09-974-300-5445	Sequence 5445, Ap
c 45	43.6	5.1	885	10	US-09-070-927A-374	Sequence 374, App

ALIGNMENTS

RESULT 1

US-10-084-205-63
; Sequence 63, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Chol, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-084-205-63

Query Match	100.08	Score 861	DB 9	Length 861
Best Local Similarity	100.08	Pred. No. 7.8e-175		
Matches 861	Conservative	0	Mismatches	0
		0	Indels	0
		0	Gaps	0
Oy	1	ATGACGAATCTACCGATCAATAAATAATAGATGAAGTCAATAATATCGGTGGG	60	
Db	1	ATGACGAATCTACCGATCAATAAATAATAGATGAAGTCAATAATATCGGTGGG	60	
Oy	61	ATAAATAATCAGTAATGGATCTACCTAGTAGAAGAAAGTATGTTGTTATTCATTAATGCT	120	
Db	61	ATAAATAATCAGTAATGGATCTACCTAGTAGAAGAAAGTATGTTGTTATTCATTAATGCT	120	
Oy	121	GGAGTAAACGATCCGACCGAGTCTGTTATTCACCTTTAGATCACTAATACCGAG	180	
Db	121	GGAGTAAACGATCCGACCGAGTCTGTTATTCACCTTTAGATCACTAATACCGAG	180	
Oy	181	TATGAGTTAGGTATGTAAGAGCGCAATTCACCTAGAAAATGATTCATACATATTCACCTATT	240	
Db	181	TATGAGTTAGGTATGTAAGAGCGCAATTCACCTAGAAAATGATTCATACATATTCACCTATT	240	

Qy	241	CATGATGACCTACGACGATGATATGATGATATCGACGAGGAAAAATTACAATCAT	300
Db	241	CATGATGACCTACGACGATGATATGATGATATCGACGAGGAAAAATTACAATCAT	300
Qy	301	AAAGTATATGGTGGAGTGCAGTATTTACGAGGTGATGCTTTATTAACTAAAGCATTT	360
Db	301	AAAGTATATGGTGGAGTGCAGTATTTACGAGGTGATGCTTTATTAACTAAAGCATTT	360
Qy	361	GAACCTATTTCACGCTGATGATAGATTAACTGATGAAGTAAAAATAAAAAGTTCTCAACGG	420
Db	361	GAACCTATTTCACGCTGATGATAGATTAACTGATGAAGTAAAAATAAAAAGTTCTCAACGG	420
Qy	421	CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGTCAAACTGTAGATGACAAAGC	480
Db	421	CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGTCAAACTGTAGATGACAAAGC	480
Qy	481	GAAGGCCAACCAATTGATCTTTGAAACTTTGGAAATGATACACAAAAACAAACACGAGCA	540
Db	481	GAAGGCCAACCAATTGATCTTTGAAACTTTGGAAATGATACACAAAAACAAACACGAGCA	540
Qy	541	TTATTAACTTTTGGCGTTATGAGTGCACGAGATATCGCTAAATGTCGATGATCAACAATAA	600
Db	541	TTATTAACTTTTGGCGTTATGAGTGCACGAGATATCGCTAAATGTCGATGATCAACAATAA	600
Qy	601	GAACATTTAGAAAGCTTAGTATTATCTAGTATGATGTTCCAGATTAAAGATGATTATA	660
Db	601	GAACATTTAGAAAGCTTAGTATTATCTAGTATGATGTTCCAGATTAAAGATGATTATA	660
Qy	661	TTAGACTGCTATGCTGATGAAGCAAAAGTTAGCTAAAAAGTGGCGAGCGATCTTCAAAAT	720
Db	661	TTAGACTGCTATGCTGATGAAGCAAAAGTTAGCTAAAAAGTGGCGAGCGATCTTCAAAAT	720
Qy	721	AATAAAAAGTACGTCAGTGCTTTATTAGGGAAAGATGGCGCAGAGAATAAATGACTTAT	780
Db	721	AATAAAAAGTACGTCAGTGCTTTATTAGGGAAAGATGGCGCAGAGAATAAATGACTTAT	780
Qy	781	CATAGACGCGCAGCAGTGGATGAACTAACGCAAAATTTGATGAACAAATTCAAATCAAAAACAC	840
Db	781	CATAGACGCGCAGCAGTGGATGAACTAACGCAAAATTTGATGAACAAATTCAAATCAAAAACAC	840
Qy	841	TTATTAGAAATCGTTGATTATA	861
Db	841	TTATTAGAAATCGTTGATTATA	861

```

RESULT 2
US-09-925-637-63
Sequence 63, Application US/09925637
Patent No. US2002010338A1
GENERAL INFORMATION:
; APPLICANT: ChOI
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925.637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-637-63

```

Query Match	100.0%;	Score 861;	DB 10;	Length 861;
Best Local Similarity	100.0%;	Pred. No. 7.8e-175;		
Matches 861;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	ATCAGCAATCTACCGGATGAATAATTAATAGATGAAGTCAATTAATCAATATTCGGTTGCG	60		
DB 1	ATCAGCAATCTACCGGATGAATAATTAATAGATGAAGTCAATTAATCAATATTCGGTTGCG	60		
QY 61	ATAAATAAATCAGTAAATGGATGCTCAGCTAGAGAAAGATGTGTGTATTCATTAAATGCT	120		
DB 61	ATAAATAAATCAGTAAATGGATGCTCAGCTAGAGAAAGATGTGTGTATTCATTAAATGCT	120		
QY 121	GGAGGTAACCGCATCCGACCGAGTCTGTATTACTCAGTAAATGATTCATACATATTCACCTAAT	180		
DB 121	GGAGGTAACCGCATCCGACCGAGTCTGTATTACTCAGTAAATGATTCATACATATTCACCTAAT	180		
QY 181	TATGAGTTAGGTATGAAGAGCGCAATGGCAGTAAATGATTCATACATATTCACCTAAT	240		
DB 181	TATGAGTTAGGTATGAAGAGCGCAATGGCAGTAAATGATTCATACATATTCACCTAAT	240		
QY 241	CATGATGACCTACCGGATGATGATGATATCGACGAGGAAATTAACAAATCAT	300		
DB 241	CATGATGACCTACCGGATGATGATGATATCGACGAGGAAATTAACAAATCAT	300		
QY 301	AAAGTATATGGTGGAGTGGCATATAGCAGGTGATGCTTTTAACTAAAGCATTT	360		
DB 301	AAAGTATATGGTGGAGTGGCATATAGCAGGTGATGCTTTTAACTAAAGCATTT	360		
QY 361	GAACCTATTTCAGTGGATGATGATTAAGTGAAGTAAATATAAGTCTTACAACGG	420		
DB 361	GAACCTATTTCAGTGGATGATGATTAAGTGAAGTAAATATAAGTCTTACAACGG	420		
QY 421	CTGTCAATAGCAAGTGGTCAATGTTGGAATGGTGGCGGTCAATGTTAGATATGCAACG	480		
DB 421	CTGTCAATAGCAAGTGGTCAATGTTGGAATGGTGGCGGTCAATGTTAGATATGCAACG	480		
QY 481	GAAGGCCAACCAATTGATCTTGAACCTTTGGAATGATACACAAAACAAAACAGGAGCA	540		
DB 481	GAAGGCCAACCAATTGATCTTGAACCTTTGGAATGATACACAAAACAAAACAGGAGCA	540		
QY 541	TTATTAACTTTTCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATCAACTAAA	600		
DB 541	TTATTAACTTTTCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATCAACTAAA	600		
QY 601	GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAGATGATTTA	660		
DB 601	GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAGATGATTTA	660		
QY 661	TTAGACTGCTATGGTGGATGAAGCAAGTTAGGTAAAAAGTGGCAGCGATCTTGAAT	720		
DB 661	TTAGACTGCTATGGTGGATGAAGCAAGTTAGGTAAAAAGTGGCAGCGATCTTGAAT	720		
QY 721	AATAAAGTACGTACGTGAGTTTATAGGAAAGATGGCGCAGAGATAAATTGACTTAT	780		
DB 721	AATAAAGTACGTACGTGAGTTTATAGGAAAGATGGCGCAGAGATAAATTGACTTAT	780		
QY 781	CATAGAGCGCAGCAGTGGATGAACTAAGCAAAATGATGAACAATTAATCAATCAAAACAC	840		
DB 781	CATAGAGCGCAGCAGTGGATGAACTAAGCAAAATGATGAACAATTAATCAATCAAAACAC	840		
QY 841	TTATTAGAAATCGTTGATTTA 861			
DB 841	TTATTAGAAATCGTTGATTTA 861			

RESULT 3

US-08-781-986A-155/c

: Sequence 155, Application US/08781986A

: Publication No. US20030054436A1

: GENERAL INFORMATION:

: APPLICANT: Charles Kunsch

: TITLE OF INVENTION: staphylococcus aureus Polynucleotides and Sequences

: NUMBER OF SEQUENCES: 5255

```
;;
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Human Genome Sciences, Inc.
;; STREET: 9410 Key West Avenue
;; CITY: Rockville
;; STATE: Maryland
;; COUNTRY: USA
;; ZIP: 20850
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;; COMPUTER: HP Vectra 486/33
;; OPERATING SYSTEM: MSDOS version 6.2
;; SOFTWARE: ASCII Text
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/781,986A
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Benson, Bob
;; REGISTRATION NUMBER: 30,446
;; REFERENCE/DOCKET NUMBER: PB248pp
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;;
;; INFORMATION FOR SEQ ID NO: 155:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1893 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;;
;; US-08-781-986A-155
;;
;; Query Match 100.0%; Score 861; DB 7; Length 1893;
;; Best Local Similarity 100.0%; Pred. No. 1e-174;
;; Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;;
Qy 1 ATGACGAATCTACCGATGAATAAATAATAGATGAAGTCAATAATGAATTCGTTGCG 60
Db 1314 ATGACGAATCTACCGATGAATAAATAATAGATGAAGTCAATAATGAATTCGTTGCG 1255
Qy 61 ATAATAAATCAGTAAGTACATCAGTCAGTGAAGAAAGTATGTTGTTATTCATTAATGCT 120
Db 1254 ATAATAAATCAGTAAGTACATCAGTCAGTGAAGAAAGTATGTTGTTATTCATTAATGCT 1195
Qy 121 GGAGTAAACGCATCCGACCGAGTCTCTGTTATTAATCTACTTTAGATTCACATAAATACCGAG 180
Db 1194 GGAGTAAACGCATCCGACCGAGTCTCTGTTATTAATCTACTTTAGATTCACATAAATACCGAG 1135
Qy 181 TATGATTTAGTATGAAGAGCGCAATTCGACTAGAAATGATTCATATATTCATTTAT 240
Db 1134 TATGATTTAGTATGAAGAGCGCAATTCGACTAGAAATGATTCATATATTCATTTAT 1075
Qy 241 CATGATGACCTACCGAGCGATGAATAATGATGATTCGAGGAGAAATTAACAATCAT 300
Db 1074 CATGATGACCTACCGAGCGATGAATAATGATGATTCGAGGAGAAATTAACAATCAT 1015
Qy 301 AAAGTATATGTTGAGTGGACTGCGATATATGACGAGTGTGTTTATTAATCAAAAGCATTT 360
Db 1014 AAAGTATATGTTGAGTGGACTGCGATATATGACGAGTGTGTTTATTAATCAAAAGCATTT 955
Qy 361 GAATTTATCAAGTATGATAGATTAATCAATGATGAAGTAAATAAATAAAGTTCTACAACGG 420
Db 954 GAATTTATCAAGTATGATAGATTAATCAATGATGAAGTAAATAAATAAAGTTCTACAACGG 895
Qy 421 CTGTCAATAGCAAGTGTGATGTTGGAATGATGTCGCGGTCAAAATGTTAGATATCAAGC 480
Db 894 CTGTCAATAGCAAGTGTGATGTTGGAATGATGTCGCGGTCAAAATGTTAGATATCAAGC 835
Qy 481 GAAGGCCAACCAATTTGATCTTTGAAATGATACACAAACAAACAAACAGGAGCA 540
Db 834 GAAGGCCAACCAATTTGATCTTTGAAATGATACACAAACAAACAAACAGGAGCA 775
;;
541 TTATTAACCTTTTGGCGTTATGAGTGCACGACATATCGCTAATGTCGATGATACAACTAAA 600
774 TTATTAACCTTTTGGCGTTATGAGTGCACGACATATCGCTAATGTCGATGATACAACTAAA 715
601 GAACATTTAGAAAGTTATAGTATATCATTTAGTATGATGTTCCAGATTAAAGATGATTAA 660
714 GAACATTTAGAAAGTTATAGTATATCATTTAGTATGATGTTCCAGATTAAAGATGATTAA 655
661 TTAGACTGCTATGTTGATGAAGCAAGTTAGTAAAGTGGGAGGAGTCTTCAAAAT 720
654 TTAGACTGCTATGTTGATGAAGCAAGTTAGTAAAGTGGGAGGAGTCTTCAAAAT 595
721 AATAAAGTACGTAGCTGAGTTTATAGGAAAGATGCGGAGAGATGATTAATTTGACTTAT 780
594 AATAAAGTACGTAGCTGAGTTTATAGGAAAGATGCGGAGAGATGATTAATTTGACTTAT 535
781 CATAGAGACGAGCAGTGGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 840
534 CATAGAGACGAGCAGTGGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 475
841 TTATTAAGAAATCGTTGATTAA 861
474 TTATTAAGAAATCGTTGATTAA 454
;;
RESULT 4
US-09-815-242-8485
; Sequence 8485, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8485
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(882)
; US-09-815-242-8485
;;
Query Match 98.3%; Score 846.6; DB 10; Length 882;
Best Local Similarity 99.0%; Pred. No. 9.4e-172;
Matches 852; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

QY 1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATGAATGAATATCGGTGG 60
Db 1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATGAATGAATATCGGTGG 60
QY 61 ATAAATAAATCAGTAATGGATCTAGCTAGAGAAAGTATGTGTATTCATTAATGCT 120
Db 61 ATAAATAAATCAGTAATGGATCTAGCTAGAGAAAGTATGTGTATTCATTAATGCT 120
QY 121 GGAGGTAAGCGCATCCGACCGAGTCTGTATTACTCAGTCTAGATTCAGTAAATACCGAG 180
Db 121 GGAGGTAAGCGCATCCGACCGAGTCTGTATTACTCAGTCTAGATTCAGTAAATACCGAG 180
QY 181 TATGAGTTAGGTATGAAGCGCAATGCACTAGAAATGATTCATCATATTCACATTAT 240
Db 181 TATGAGTTAGGTATGAAGCGCAATGCACTAGAAATGATTCATCATATTCACATTAT 240
QY 241 CATGATGACCTACCGAGTGGATGATGATGATTCGACGAGGAAATTAACAAATCAT 300
Db 241 CATGATGACCTACCGAGTGGATGATGATGATTCGACGAGGAAATTAACAAATCAT 300
QY 301 AAAGTATATGGTGGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 301 AAAGTATATGGTGGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 GAATTTATTTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 GAATTTATTTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 CTGCTCAATAGCAAGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 421 CTGCTCAATAGCAAGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GAAGGCCAACCAATTCATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 481 GAAGGCCAACCAATTCATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 TTATTAACATTTTCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 541 CTATTAACATTTTCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 GAACATTTAGAAGTTATAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 GAACATTTAGAAGTTATAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TTAGACTGCTATGGTGAAGCAAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 720
Db 661 TTAGACTGCTATGGTGAAGCAAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 720
QY 721 AATAAAGTACGTACGTGAGTTTATAGGAAAGTGGCGCAGAGATGATGATGATGATGATGAT 780
Db 721 AATAAAGTACGTACGTGAGTTTATAGGAAAGTGGCGCAGAGATGATGATGATGATGATGAT 780
QY 781 CATAGAGCGCAGCAGTGGATGAATCAAGCAATTTGATGATGATGATGATGATGATGATGATGAT 840
Db 781 CATAGAGCGCAGCAGTGGATGAATCAAGCAATTTGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TTATTAAGAATCGTTGATTTA 861
Db 841 TTATTAAGAATCGTTGATTTA 861

RESULT 5

US-09-815-242-4184
; Sequence 4184, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4184
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4184

Query Match 96.6%; Score 831.6; DB 10; Length 864;
Best Local Similarity 98.9%; Pred. No. 1.9e-168;
Matches 837; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 16 ATGAATAAATTAATAGATGAAGTCAATGAATATCGGTCCGATAAATAATCAGTA 75
Db 1 ATGAATAAATTAATAGATGAAGTCAATGAATATCGGTCCGATAAATAATCAGTA 60
QY 76 ATGGATACCTAGCTAGAGAAAGTATGTTGTTATTAATTAATGCTGGAGTAACGCATC 135
Db 61 ATGGATACCTAGCTAGAGAAAGTATGTTGTTATTAATTAATGCTGGAGTAACGCATC 120
QY 136 CGACCACTGCTGTATTACTCTACTTTAGATTCTACTAAATACCGAGTATGAGTTAGTATG 195
Db 121 CGACCACTGCTGTATTACTCTACTTTAGATTCTACTAAATACCGAGTATGAGTTAGTATG 180
QY 196 AAGAGCGCAATTCGACTAGAAATGATTCATACATATTCATTTATTCATGATGACCTACCA 255
Db 181 AAGAGCGCAATTCGACTAGAAATGATTCATACATATTCATTTATTCATGATGACCTACCA 240
QY 256 GCGATGGATAATGATGATGATTCGACGAGGAAATTAACAAATCATTAAGTATATGTTAG 315
Db 241 GCGATGGATAATGATGATTCGACGAGGAAATTAACAAATCATTAAGTATATGTTAGTATG 300
QY 316 TGGACTCGCATATTACGAGTGTATGCTTTTAACTAAAGCATTTTCAAGTATTTCAAGT 375
Db 301 TGGACTCGCATATTACGAGTGTATGCTTTTAACTAAAGCATTTTCAAGTATTTCAAGT 360
QY 376 GATGATAGATTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
Db 361 GATGATAGATTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 436 GGTGATGTTGGATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
Db 421 GGTGATGTTGGATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 496 GATCTTGAACCTTTGGAAATGATACACAAACAGAGGAGCATTTATTAACCTTTTGGC 555
Db 481 GATCTTGAACCTTTGGAAATGATACACAAACAGAGGAGCATTTATTAACCTTTTGGC 540
QY 556 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGATGATGATGATGATGATGATGATGAT 615
Db 541 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGATGATGATGATGATGATGATGATGAT 600
QY 616 TATAGTTATCAATTAGTATGATGTTCCAGATTAAGATGATTTATTAGCTGCTATGTT 675

Db 601 TATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTATTATAGACTGCTATCGT 660
QY 676 GATGAAGCAAAAGTTAGGTAAAAAGTGGCGACGATCTTGAATAATAATAAAGTACGTAC 735
Db 661 GATGAAGCAAGTTAGGTAAAAAGTGGCGACGATCTTGAATAATAATAAAGTACATAC 720
QY 736 GTGAGTTTATTAGGAAAGATGGCGCAGCAAGATTAATTTGACTTATCATAGAGACGACGA 795
Db 721 GTGAGTTTATTAGGAAAGATGGCGCAGCAAGATTAATTTGACTTATCATAGAGACGACGA 780
QY 796 GTGAGTGAACCTAACCAAAATGATGAACAATTCATACAAACACCTTTATTGAATCGTT 855
Db 781 GTGAGTGAACCTAACCAAAATGATGAACAATTCATACAAACACCTTTATTGAATCGTT 840
QY 856 GATTATA 861
Db 841 GATTATA 846

RESULT 6

US-09-815-242-2822/c

; Sequence 2822, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2822

; LENGTH: 413

; TYPE: DNA

; ORGANISM: Staphylococcus aureus

US-09-815-242-2822

Query Match 48.0%; Score 413; DB 10; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.6e-79;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 TTAGATTCACATAAATACCGAGTATGAGTTAGGTAGATGAAGCGCAATTCGACTAGAAATG 219
Db 413 TTAGATTCACATAAATACCGAGTATGAGTTAGGTAGATGAAGCGCAATTCGACTAGAAATG 354
QY 220 ATTATACATATTCACCTTATTCATGATGACCTACCGCGATGGATTAATGATGATTCGGA 279
Db 353 ATTATACATATTCACCTTATTCATGATGACCTACCGCGATGGATTAATGATGATTCGGA 294
QY 280 CGAGGAAATTAACAAATCATTAAGTATATGGTGGTGGACTGCGATATTAGCAGGTGAT 339

Db 293 CGAGGAAATTAACAAATCATTAAGTATATGGTGGTGGACTGCGATATTAGCAGGTGAT 234
QY 340 GCTTTTATTAACATAAAGCATTTTGAACCTTATTTCAAGTGATGATAGATTAACCTGATGAAGTA 399
Db 233 GCTTTTATTAACATAAAGCATTTTGAACCTTATTTCAAGTGATGATAGATTAACCTGATGAAGTA 174
QY 400 AAAATAAAAGTTCTTACAACGGCTGTCAATAGCAAGTGTGATGTTGGAAATGTCGCGCGT 459
Db 173 AAAATAAAAGTTCTTACAACGGCTGTCAATAGCAAGTGTGATGTTGGAAATGTCGCGCGT 114
QY 460 CAAATGTTAGATATCAAAAGCGCAAGCCCAACCAATTTGATCTTGAACCTTTGGAATGATA 519
Db 113 CAAATGTTAGATATCAAAAGCGCAAGCCCAACCAATTTGATCTTGAACCTTTGGAATGATA 54
QY 520 CACAAAACAAAACAGGAGCATTATTAACTTTTTCGGTTTATGAGTGCAGCAGA 572
Db 53 CACAAAACAAAACAGGAGCATTATTAACTTTTTCGGTTTATGAGTGCAGCAGA 1

RESULT 7

US-09-815-242-3275/c

; Sequence 3275, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3275

; LENGTH: 337

; TYPE: DNA

; ORGANISM: Staphylococcus aureus

US-09-815-242-3275

Query Match 39.1%; Score 337; DB 10; Length 337;
Best Local Similarity 100.0%; Pred. No. 6e-63;
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 TACATATTCACCTTATTCATGATGACCTACCGCGATGATTAATGATGATTAATTCGAGG 284
Db 337 TACATATTCACCTTATTCATGATGACCTACCGCGATGATTAATGATGATTAATTCGAGG 278
QY 285 AAAATTAACAATCAATAAGTATATGGTGGTGGACTGCGATATTAGCAGGTGATGCTTT 344
Db 277 AAAATTAACAATCAATAAGTATATGGTGGTGGACTGCGATATTAGCAGGTGATGCTTT 218
QY 345 ATTAACCTAAGCATTGGAACCTTATTCAAGTATGATGATTAATGATGATGAAGTAAAT 404

Db 217 ATTAACTAAAGCATTTGAACCTTATTTCAAGTGATGATAGATTAACTGATCAAGTAAATAAT 158
QY 405 AAAAGTCTTACACGCGCTGTCAATAGCAAGTGTGTGGAATGTCGGCGGTCAAAAT 464
Db 157 AAAAGTCTTACACGCGCTGTCAATAGCAAGTGTGTGGAATGTCGGCGGTCAAAAT 98
QY 465 GTTAGATATGCAAGCGAAGCGCAACCAATTGATCTTGAACCTTTGGAATGATACACAA 524
Db 97 GTTAGATATGCAAGCGAAGCGCAACCAATTGATCTTGAACCTTTGGAATGATACACAA 38
QY 525 AAAAAAAGCAGGAGCATTTTAACTTTTTCGGGTATG 561
Db 37 AAAAAAAGCAGGAGCATTTTAACTTTTTCGGGTATG 1

RESULT 8

US-09-815-242-6533

Sequence 6533, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA-011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6533

LENGTH: 882

TYPE: DNA

ORGANISM: Enterococcus faecalis

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(882)

US-09-815-242-6533

Query Match

Best Local Similarity 22.9%; Score 196.8; DB 10; Length 882;

Matches 413; Conservativity 0; Mismatches 317; Indels 6; Gaps 2;

QY 92 AAGAAGTATGTGTATTCAATTAATGCTGGAGGTAAGCGCATCCAGCAGTCTCTGTAT 151
Db 89 AGAAGCATGTCTACTCAATACAGCGAGGTGGCAAGCATACGCCGCTATTAGTCC 148
QY 152 TACTCACTTTAGATTCACATAACCGAGTATCA---GTTAGTATGAAGACGCAATG 208
Db 149 TAACAACAGTGGCGCTTTTCAAAAGAGATGGAACGCAAGCATATCAAGTGGCTCT 208
QY 209 CACTAGAATGATTCATACATATTCCTATTTCATGATGACCTACCAGCGGATGATG 268

Db 209 CTTTAGAGATGATTCATACGTATTTCATTTAATTCATGATGATTTACCAGCAATGGATGATG 268
QY 269 ATGATTATCGAGGAGGAAATTAACAAATATATTAAGTATATGTTAGTGTGAGTGCAGTAT 328
Db 269 ATGATTATCGAGGAGGAAATTAACAAATATATTAAGTATATGTTAGTGTGAGTGCAGTAT 328
QY 329 TAGCAGGTGATGCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 388
Db 329 TAGCTGGGCGGCTTATTAAACAGGCGCATTTTCAGTTACTTTC---TTTGAGCCAATTAG 385
QY 389 CTGATCAAGTAAATAATAAAGTCTTACAACGGCTGTCAATAGCAAGTGTGATGTTGGAA 448
Db 386 GCTTAAGTAAAAAGTTTTCATGATGCAACAACTGGCAAAAGCGCGGGAATCAAGGCA 445
QY 449 TGGTCGGCGGTCAAAATGTTAGATATGCAAAAGCGGCAACCAATTTGATCTTTGAAACTT 508
Db 446 TGGTTCGGCGGCAAAATGTTAGATATGCAAAAGCGGCAACCAATTTGATCTTTGAAACTT 505
QY 509 TGGAAATGATACACAAAAACAGGAGCATTTTAACTTTTTCGGGTTATGAGTGCAG 568
Db 506 TAGCGGCTGTTTCATGAAAAAGAAACCGGAGCGCTAATTGAATTCGCTTAAATTGCTGGG 565
QY 569 CAGATATCGCTAATGCTGATGATACAACTAAAGAACATTTAGNAAGTTATAGTTATCATT 628
Db 566 GCGTATTAGCAATCAACAGAGAAAGTCAATGTTTACTGACACAGTTCGCGCATCACT 625
QY 629 TAGGTATGATGTTCCAGATTTAAAGATGATTTTATAGACTGCTATGTTGATGAAGCAAGT 688
Db 626 ATGGCTTGGCTTTTCAAAATTCGTCGATGATGATTTAGTATGCCACAGTACAGAGCAGATT 685
QY 689 TAGGTAAAAAGTGGCGAGCGATCTTTGAAAAATAATAAAGTACGTACGTGAGTTTATTAG 748
Db 686 TAGGTAAAAAGTGGCGAGATGAAGCGTGAATAAAAAGTACGTATCCAGCTCTTCTTAG 745
QY 749 GGAAGATGGCGAGAGATTAATTCATCTATCATAGACAGCGCAGTGGATGAAGTAA 808
Db 746 GGATGCTGGCGGAAAGATGCGCTAACACATCAATTAGCAGAAGCGCGCTGTTTATTAG 805
QY 809 CGCAATGATGAACA 824
Db 806 AAAAATCAAGCAAA 821

RESULT 9

US-09-070-927A-55

Sequence 55, Application US/09070927A

Patent No. US20020120116A1

GENERAL INFORMATION:

APPLICANT: Charles A. Kunsch

Steven Barash

Patrick J. Dillon

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 982

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927A

FILING DATE: 04-May-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/046,655

FILING DATE: 1997-05-16

APPLICATION NUMBER: 60/044,031

;; FILING DATE: 1997-05-06
;; APPLICATION NUMBER: 60/066,009
;; FILING DATE: 1997-11-14
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kenley K. Hoover
;; REGISTRATION NUMBER: 40,302
;; REFERENCE/DOCKET NUMBER: PB369
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 55:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7528 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-070-927A-55

Query Match 22.6%; Score 195; DB 10; Length 7528;
Best Local Similarity 56.9%; Pred. No. 4.3e-32;
Matches 398; Conservative 0; Mismatches 295; Indels 16; Gaps 2;
Qy 92 AAGAAAGTATGCTGCTATTAATGCTGGAGTAAACGCATCCGACCGAGTCTGTAT 151
Db 6730 AAGAGCAATGCTCTACTCATACGCGAGTGGCAACGATTACGCCGCTATTAGTG 6789
Qy 152 TACTCATTAGTTCATTAATACCGAGTAGTA---GTTAGGTATGAAGAGCCCAATTG 208
Db 6790 TAACAACAGTGGCGCTTTTCAAAAAGAGATGGAACACAGACTATCAAGTGGCTGCT 6849
Qy 209 CACTAGAAATGATTCATACATATTCACCTTATTCATGATGCTACCAGCGATGATATG 268
Db 6850 CTTTAGAGATGATTCATACGCTATTCATTAATTCATGATGATTTACCAGCAATGACGATG 6909
Qy 269 ATGATTATCGAGGAGGAAATTAACAATCATATAAGTATATGTTGAGTGGACTCGATAT 328
Db 6910 ATGATTACGTCGTGGCAACCAATCATTAAGTCTTTGGTGAAGCGACTGCCATT 6969
Qy 329 TAGCAGGTGATGCTTTATTAATTAAGTATTAAGCTTTTAAAGTATGATGATGATTA 388
Db 6970 TAGCAGGGGAGCGCTTATTAAACAGTGCATTTCAAGTTCAGTTCTTTC---TTTGAAGTCAATAG 7026
Qy 389 CTGATGAAGTAAATAAAGTCTTACAAACGCTGTCAATAGCAAGTGTGATGTTGAA 448
Db 7027 GCTTAAGTGAAGTCTTTTACTGATGCAACAACTGGCAAGGCAAGCGGGGAATCAAGCA 7086
Qy 449 TGGTCGGCGTCAATGTTAGATATGCAAGGAGGCAACCAATGATCTTGAAGCTT 508
Db 7087 TGGTTCCGGCAATGGGTGATATGAAGGAGAAAGTGCAGCTTGACTTTAGAGAGT 7146
Qy 509 TGGAAATGATACACAAACAAACAGAGCAATTAATTAACCTTTTGGGTTATGATGCGAG 568
Db 7147 TAGCGGCTGTTCATGAAGAAACCGGAGCGCTAATTAAGTTCGCTTAATTTGCTGGG 7206
Qy 569 CAGATATGCTGAATGCTGATACAACTAAAGAACATTTAGAAAGTTATAGTATCATTT 628
Db 7207 CGGTATTAGCCCAATCAACAGAAAGTCAATGTTTACTGACACAGTTTCGGCTCACT 7266
Qy 629 TAGGTATGATCTCCAGATTAAAGATGATTTATTAGACTGCTATGTTGATGAAGCAAGT 688
Db 7267 ATGGCTTGGCTTTTCAATTCGTGATGACTTATTAGTGCACAAAGTACAGAGCAGATT 7326
Qy 689 TAGGTAAAGAGTGGGAGCGCTTTGAAATAATAAAGTACGTACGTAGTTATTATTA 748
Db 7327 TAGGTAAAGAGTAGGCGGAGATGAAGCGTTGAATAAAGTACGTATCCAGCTCTCTAG 7386
Qy 749 GGAAGATGGCGCAGAGATAAATGACTTATCATAGAG 787
Db 7387 GGAATGCTGGCGGAAAGATCGCTAAACATCAATTAG 7425

RESULT 10

US-09-974-300-1015
;; Sequence 1015, Application US/09974300
;; Patent No. US20020146721A1
;; GENERAL INFORMATION:
;; APPLICANT: Berka, Randy M.
;; APPLICANT: Clausen, Ib Groth
;; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;; TITLE OF INVENTION: Expression
;; FILE REFERENCE: 10085.500-US
;; CURRENT APPLICATION NUMBER: US/09/974,300
;; CURRENT FILING DATE: 2001-10-05
;; PRIOR APPLICATION NUMBER: 09/680,598
;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: 60/279,526
;; PRIOR FILING DATE: 2001-03-27
;; NUMBER OF SEQ ID NOS: 8481
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1015
;; LENGTH: 811
;; TYPE: DNA
;; ORGANISM: Bacillus licheniformis
US-09-974-300-1015

Query Match 19.0%; Score 164; DB 10; Length 811;
Best Local Similarity 52.9%; Pred. No. 8e-26;
Matches 378; Conservative 0; Mismatches 330; Indels 6; Gaps 1;
Qy 88 CTAGAAGAAAGTATGTTGTTTCAATTAATGCTGGAGTAAACGCATCCGACCGAGTCTGTG 147
Db 97 CTCAAGGAATCCATGCTGCTACTCATTTGGAAGCGGGGGGAAACGGCTGAGACCGGATTTG 156
Qy 148 TTATTACTCACTTTAGATTCTACTTAATACCGAGTATGAGTTAGGTATGAAGAGCCCAATT 207
Db 157 GTTCTTGGCGTCTTTCATCGTCGCGTAAAGAGCAAGGAAAGCGGAAATTCAGTCCGATGC 216
Qy 208 GCCTAGAAATGATTATCATATATTCATTTATTCATGATGACCTACCAGCGATGATAT 267
Db 217 CGGTTGAATGATCCATGCTATTCATTAATCCATGACGCTCCCTTGTATGATGAC 276
Qy 268 GATGATTATCGACGAGGAAATTAACAAATCATAAATATATGTTGAGTGGACTCGGATA 327
Db 277 GATGACCTCAGAGAGGTAAGCGCAACCAACCAAAATTTACGGGAGGCGAGCGCCATT 336
Qy 328 TTACGAGTGTGCTTTTATTAACTAAAGCAATTTGACATTTTTC-----AGTGATGAT 381
Db 337 CTCGGGAGAGCGCTTATTAAACCGAAAGCTTCAAAATGATCACTTCCAATATGCTTCT 396
Qy 382 AGATTAACTGATGAAGTAAATAATAAAGTCTTACAAACGCTGTCAATAGCAAGTGTGAT 441
Db 397 GATGATCCGCTGAAGAGCGATCAGACTGGTGAACGAGCTGATTTCCGACGCGGGGCC 456
Qy 442 GTTGGAAATGTCGGCGGTCAAAATGTTAGATATGCAAAAGCGAAGCGCAACCAATTTGATCT 501
Db 457 GAAGGCAATGTCGGCGGTGAGATTTTATGATATGGAAGCGGAATCAAGTCTGTATCTCT 516
Qy 502 GAACTTTGGAATGATACACAAACAAACAGAGGAGTATTAACTTTTTCGGGTTATG 561
Db 517 GATGAACCTGACGCGAATCCACGAGGAAACCGCTAAACCTCTCAGCTTTCAGCGCTATT 576
Qy 562 AGTCGACGAGATATCGCTAAATGTCGATGATCAACTAAAGAACATTTAGAAAGTTATAGT 621
Db 577 CGGGAGCGCATTTGAGATGATGCTGGAAGAGATCGAAAGCTGGGGAATTCAGC 636
Qy 622 TATCATTTAGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGTTGGTATGAA 681
Db 637 CATCATATCGCATCGGCTTTTCAGATCAGAGATGACATTTTGGACCTTTGAAGGCTCGGAG 696
Qy 682 GCAAGATTAGTAAAGAGTGGGAGCGATCTTGAATAATAAAGTACGTACGTAGT 741
Db 697 GATAAGATCGGCAACGAGTCCGATCCGAGCTTCAAAACGGAAGTTCAGCTATCCGTCG 756
Qy 742 TTATTAGGAAAGATGGCGCAGAGATAAATTTGACTTATCATAGAGCGACGA 795

Db 757 CTCCTTCGCTGAGGGGCCCAACAACTTGTGATGACATATGAAAAAGCA 810

RESULT 11

US-09-815-242-7142

Sequence 7142, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7142

LENGTH: 888

TYPE: DNA

ORGANISM: Haemophilus influenzae

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(888)

US-09-815-242-7142

Query Match 15.3%; Score 131.4; DB 10; Length 888;

Best Local Similarity 49.9%; Pred. No. 7.8e-19;

atches 361; Conservative 0; Mismatches 356; Indels 6; Gaps 1;

QY 100 ATGTTGATTAATTAATGCTGGAGTAAACGCATCCGACCACTGCTGTATTACTCACT 159

Db 109 ATGAATACGCATTAATTAATGCTGGAGTAAACGCATCCGACCACTGCTGTATTACTCACT 168

QY 160 TTAGATTCAATAATACCAAGTATGATGTTAGTGAAGAGCGCAATTCCTACATAGATG 219

Db 169 GTCAATGCTTGGCGACAGAAACCAACCTTAGATTACGCTGCTGCCATTTGAAGCC 228

QY 220 ATTCAATATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 279

Db 229 ATTCAGGCTATTCCTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 288

QY 280 CGAGGAAATTAACAAATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 339

Db 289 CGTGACATCTTACTGTATATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 348

QY 340 GCTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 399

Db 349 GCACGCTAAAGTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 408

QY 400 AAAAATAAAGTTCTACAAGCGGTGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 459

Db 409 AAATGGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 468

QY 460 CAAATGTTAGATATGCAAGCGAAGCCCAACCAATTTGATCTTGAACCTTTTGGAAATGATA 519

Db 469 CAAAGTTTAGATCTTATTTCTGAGCATAAACAGATTTAGTTAGTGAATTAGAATTAAT 528

QY 520 CACAAACAAACAGAGCATTATTAACCTTTTGGCGTTATGAGTCAGCAGATATCGC- 578

Db 529 CATCGTAACAAAACGGGTGCTTAAATTTGGCGCATTTGAAATTTAGTTTCTATTCTTCT 588

QY 579 -----TAATGTCGATGATACAACTAAAGAACATTTTGAAGATTTATAGTTATCTATTAGGT 633

Db 589 CCGCATTTTACCGACAAAGGTTAGAACAAATCCTTAACACAAATATCCGAAGCCATTGGT 648

QY 634 ATGATGTTCCAGATTAAGATGATTTTATAGATGCTGCTATGTTGATCAAGCAAAAGTTAGGT 693

Db 649 TTAGCCTTTCAAGTTCAAGCAGATATTTAGATATTTGAAGGCGATAGTCAGAAATTTGGC 708

QY 694 AAAAAGTGGCAGCATCTTGAATAATAAAGTACGTGATGTTAGTTTATTTAGGAAA 753

Db 709 AAACAGTGGTGGCGCATCTTGAATTTAGATAAAAGTACATATCCAAAATTTACTTGGATTA 768

QY 754 GATGGCGCAGAGATAAATTTGACTTATCATAGAGCGCAGCAGTGGATGACATACGCA 813

Db 769 AGTGGTGCAAAACAAAGCGCAAGATCTATATCAAAAGTGGCTGTCTGAATTTAGAAA 828

QY 814 ATT 816

Db 829 ATT 831

RESULT 12

US-09-815-242-9175

Sequence 9175, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9175

LENGTH: 876

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(876)

US-09-815-242-9175

Query Match	14.5%	Score 124.8	DB 10	Length 876
Best Local Similarity	51.2%	Pred. No. 2e-17		
Matches 344	Conservative 0	Mismatches 322	Indels 6	Gaps 2
QY	89	TAGAAGAAGATGTTGTATTTCATTTAAATCGTGAGGTAAACGCATCCGACACAGTCTCTGT	148	
Db				
QY	80	TACGGGAGTCTGTTCTCTATTCTATTCTATGCTGGTGGCAAGGTATTCGGCCTCTTCTCT	139	
Db				
QY	149	TATTTACTCCTTTTAGATTTCACATAAA---TACCAGGATGATGAGTTAGGTATGAAGAGCGCAA	205	
Db				
QY	140	TGTTAGAAGTCTTGAAGCCTTGCAGGTGCCATCAAACTGCTCACGCGCAGGTAGCTA	199	
Db				
QY	206	TTGCACTAGAATAAGTATTCATACATATTCTATGATGACCTACACAGGATGGGATA	265	
Db				
QY	200	CTGCGCTTGGAGATGATTCATACAGGGAGCTTGATTCACGATACCTTCTCGTCTGGATG	259	
Db				
QY	266	ATGATGATTATCGCAGGAGGAAATTAACAATCATAAAGTATATGCTGGAGTGGACTGCGA	325	
Db				
QY	260	ATGACGATATTCGAAGAGGGCGTTAACCAATCACAGAAATTCGGGTGAAGCTATGGCCA	319	
Db				
QY	326	TATTAGCAGGTGATGCTTTTATTAACTAAAGCAATTTGAACCTATTATTTCAAAGTGATGATGAT	385	
Db				
QY	320	TWTTGGCTGGAGATGCGCTTATTCTTAGACCCCATATGCCTTGATTCGCGAGGCAGAT--T	376	
Db				
QY	386	TAACGTGATGAAGTAAATAAAGTCTTACAACGGCTGTCAATAGCAAGTGTGCTATGTTG	445	
Db				
QY	377	TGCCAAGTCAGATCAAGGTGGACTTGATTCGCCAACTTATCCCTTGCTTCAGGTAGTCTGG	436	
Db				
QY	446	GAATGTCGGCGGTCAATGTTAGATATGCAAGCGGCAAGCCCAACCAATTGATCTTGAA	505	
Db				
QY	437	GTATGCTGGCAGGCGAAGTTTGTGATATGAGGCGCAACACAGCACTTGTCTCTGGAA	496	
Db				
QY	506	CTTTGGAAATGATACACAAAACAAACAGGAGCATTTATTAACCTTTTGGGTTATGAGTG	565	
Db				
QY	497	AACTTCAGACTATTCATCGCCCAATAAGACTGGGAAGTTACTAGCCTATCCCTCCAGCG	556	
Db				
QY	566	CAGCAGATATCGCTAAATGTCGATGATACACTAAGAACATTTAGAAAAGTTATAGTTATC	625	
Db				
QY	557	CAGCTATTATAGCTGAATTTGTCACCTCAAAATGCAAGGTGAAGCTGAAAACCTGTGGTGAAT	616	
Db				
QY	626	ATTTAGGTATGATGTTCCAGATTAAGATGTTTATTAGACTGCTATGCTGATGAAGCAA	685	
Db				
QY	617	TGATTTGGACTTGTCTTTCAAGTCAGATGATGATCTGGATGTGACAGCTAGTTTGTAGG	676	
Db				
QY	686	AGTTAGGTAAAAAGTGGCGCAGCATCTTGAAAATAATAAAAGTACGTACGTAGTTTAT	745	
Db				
QY	677	AAATCGGTAAACACCTCAAAAGGATCTGCGAGGCAGAAAATCAACCTATCTCGCTCTGT	736	
Db				
QY	746	TAGGAAAAGATG	757	
Db				
QY	737	TGGCGCTTGGAA	748	
Db				

RESULT 13

```

RESOLUTION 13
US-09-815-242-9499
Sequence 9499, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
PROKARYOTES
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21

```

	: PRIOR APPLICATION NUMBER: 60/206,848	
	: PRIOR FILING DATE: 2000-05-23	
	: PRIOR APPLICATION NUMBER: 60/207,727	
	: PRIOR FILING DATE: 2000-05-26	
	: PRIOR APPLICATION NUMBER: 60/242,578	
	: PRIOR FILING DATE: 2000-10-23	
	: PRIOR APPLICATION NUMBER: 60/253,625	
	: PRIOR FILING DATE: 2000-11-27	
	: PRIOR APPLICATION NUMBER: 60/257,931	
	: PRIOR FILING DATE: 2000-12-22	
	: PRIOR APPLICATION NUMBER: 60/269,308	
	: PRIOR FILING DATE: 2001-02-16	
	: NUMBER OF SEQ ID NOS: 14110	
	: SOFTWARE: FastSeq for Windows Version 4.0	
	: SEQ ID NO 9499	
	: LENGTH: 876	
	: TYPE: DNA	
	: ORGANISM: Streptococcus pneumoniae	
	: FEATURE:	
	: NAME/KEY: CDS	
	: LOCATION: (1)...(876)	
	: US-09-815-242-9499	

	Query Match	14.48;	Score 124.2;	DB 10;	Length 876;
	Best Local Similarity	51.6%;	Pred. No. 2.7e-17;	Mismatches 335;	Conservative 0;
QY	89	TAGAAGAAAGTATCTTGTTATTCAATAAATGCTGGAGCTAAACGCATCCGACCAGTTCTGT	148		
DB	80	TACGGGAGTCGTCTCTATTCTATTCTATGCTGTGGCAACGGTAATTCGGCCCTTTCTCT	139		
QY	149	TATTACTCACCTTTAGATTTCACAAA---TACCAGTAGTAGTTAGGTATGAAGAGCGCAA	205		
DB	140	TGTTAGAAGTTCTGGAAGCCTTGACAGTTACCATCAAACCTGCTCACGCCGACGGTAGCTA	199		
QY	206	TTGCACATAGAAATGATTCTATACATATTTCATTTATCATGATGACCTTACCAGCGATGGATA	265		
DB	200	CTGCCTTGGAGATGATTCATACAGGGAGCTTGATTTCACGATGACCTTCTCTGCTATGGATG	259		
QY	266	ATGATGTTATTTCGACGAGGAAAAATTAACAATCATAAAGTATATGTTGAGTGGACTCGCA	325		
DB	260	ATGACGATTTCGAAGAGGGGGTTTACCAATCAACAAGAAATTCGGTGAAGCTATGGCCA	319		
QY	326	TATTACAGGTGATGCTTTATTAACATAAGCATTTGAACCTATTTCAGGTGATGATAGAT	385		
DB	320	TTTTGGCTGGAGATGCCTTATCTTAGACCATATGCTTGTATTCGCCAGGCAGAT--T	376		
QY	386	TAACTGATGAAGTAAAAATAAAGTTCTACACGGCTGTCOAATAGCAAGTGGTCAATGTTG	445		
DB	377	TGCCAAGTTCAGATTAAAGTGGACATTGATTCGCAACTTATCCCTTGTCTTCAAGTAGTCTGG	436		
QY	446	GAATGGTTCGGGGTCAATGTTAGATATGCAAGCGAAGGCCAACCAATTTGATCTTCAAA	505		
DB	437	GTAATGGTGGCAGGCGAGTTTGGATATGAGGGCGAACACACAGCACTTGTCTTTGGGAG	496		
QY	506	CTTTGGAAATGATACACAAAACAAAAACAGAGCATTATTAACTTTTTCGGTGTATGAGTG	565		
DB	497	AAC TTCAGATTATTTCATGCCAATAAGACTGGGAAGTTACTAGCCTATCCCTTCCAACGG	556		
QY	566	CAGCAGATATCGCTTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATC	625		
DB	557	CAGCTATTATAGCTGAATTGTCCTCCTGAAATGCAAGTGAAGCTGAAAACCTGTGGGTGAAT	616		
QY	626	ATTTAGGTATGATCTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGTGATCAAGCAA	685		
DB	617	TGATTGGACTTGCTTTTCAAGTCAGATGATGTACTGGATGTGACAGCTAGTTTTTGAGG	676		
QY	686	AGTTAGTAAAAAAGTGGCAGCGATCTTGAATAATAATAAAGTACCTA	734		
DB	677	AAATCGGCAAGACACCTCAAAGGATCTGCAGGCAGAAAAATCAACCTA	725		

RESULT 14

US-09-790-988-1/c

; Sequence 1, Application US/09790988

; Patent No. US20020127687A1

; GENERAL INFORMATION:

; APPLICANT: SHIGENOBU, SHUJI

; APPLICANT: WATANABE, HIDEMI

; APPLICANT: HATTORI, MASAHIRA

; APPLICANT: SAKAKI, YOSHIYUKI

; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS

; FILE REFERENCE: 081356/0159

; CURRENT APPLICATION NUMBER: US/09/790,988

; CURRENT FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: JP2000-107160

; PRIOR FILING DATE: 2000-04-07

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 640681

; TYPE: DNA

; ORGANISM: Buchnera sp.

; 09-790-988-1

Query Match 13.3%; Score 114.2; DB 10; Length 640681;

Best Local Similarity 51.4%; Pred. No. 4.2e-14;

Matches 319; Conservative 0; Mismatches 293; Indels 9; Gaps 2;

QY 88 CTAGAGAAAGATGTTGTATTCATTAAATGCTGGAGGTAAACGCGATCGACCGAGTCTCG 147

Db 511231 CTTTTAAAGCAATGAATAGTAGTATTTTCAGGTAGTAAAGATTCGTTTCATCTTTA 511172

QY 148 TTATTACTACATTGATTCACCTAAATACGGAGTAGTATGAGTATGAGAGCGCAATT 207

Db 511171 ATATATCAACTGGAGATGTTTTTAAAGTTAATATTACTACATTAGATGTATATCTACT 511112

QY 208 GCACATAGATGATTCATACATATTCATGATGACCTACCGAGCGATGATAAT 267

Db 511111 GCATCGAATTTATTCATCATACTCTTTAATACATGATGACTACCTGTATGGTAAT 511052

QY 268 GATGATTATCGAGAGAGAAATTAACAAATCATATAATGATGATGAGTGGAGTGGGATA 327

Db 511051 GATAAATTTAGAGAGAGAAAGATATCTGTCAATGTAATAATACGTTCAAAAGTACTTCTTTA 510992

QY 328 TTAGCAGGTGATGCTTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGAT 381

Db 510991 CTTGCTGGTGTGCTTTGCAAGTCTTGGCTTTAATATTCTATCAACAGTTTTTATGCA 510932

QY 382 AGATTAACTGATGAAGTAAATAAATAAAGTTCTACAACGCTGTCAATAGCAAGTGTGAT 441

Db 510931 AATGATCTAATTTAAACGATATAAATAATGATTTCTGAATATCTTATTCATTTGTTCA 510872

QY 442 GTTGGAAATGTCGCGCGTCAATGTTAGATGCAAGGCAAGGCAACCAATTTGATCTT 501

Db 510871 TCTGGGATGTCATGCTCAAAATTTAGATTTAGAACGAGAAAAAAGATGTAAATTTA 510812

QY 502 GAAACTTTGGAATGATACACAAACAAACAGGAGCATTTAATCTTTTGGGTATG 561

Db 510811 TCTGAATTAAGAATAATTAATTTATATAAACTCTCTTTTAAATGCTGCTGCTGCTGCT 510752

QY 562 AGTGACGAGATGCTCAATGTCGATGATACAACTA---AAGAACATTTAGAAAGTTAT 618

Db 510751 TTAGTATATTTTCTCTAATAATTTTCTAATCTAATATTAATATTAATATTAATCTTTT 510692

QY 619 AGTTATCATTTAGGTATGATGTTTCCAGATTAAGATGATTTATTAAGACTGCTATGGTAT 678

Db 510691 TCGATTTCTATTGGTTTGGCATTTCAAAATTCAGGATGATATTTTAGATTTTAAAAAAGAT 510632

QY 679 GAAGCAAGTTAGTAAAAA 699

Db 510631 AGTGTAAAAACAGACAAATAA 510611

RESULT 15

US-09-815-242-5972

; Sequence 5972, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5972

; LENGTH: 900

; TYPE: DNA

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(900)

; US-09-815-242-5972

Query Match

Best Local Similarity 11.9%; Score 102.4; DB 10; Length 900;

Matches 363; Conservative 0; Mismatches 386; Indels 9; Gaps 2;

QY 78 GGATACCTAGCTAGAGAAAGATGTTGTATTCATTAAATGCTGGAGGTAAACGCGATCG 137

Db 84 GAACACTCCGGTGGTGAACCATGCGATGCGCATTTATTAGTGGTAGCGGCTCGG 143

QY 138 ACCAGTCTGTATTACTACTTTAGATTCACCTAAATACCGAGTATGAGTTAGGTATGAA 197

Db 144 ACCTTTCTGTTTATGCCACCGGTCATATGTTGCGGTTAGCACAAACACGCTGGACGC 203

QY 198 GAGCGCAATGTCACCTAGAAATGATTCATACATATTCACCTATTCATGATGACCTACCAGC 257

Db 204 ACCCGCTGCGCGGTGAGTGTATCCACGCTTACTACTTAATTCATGATGATTTACCGGC 263

QY 258 GATGCAATATGATGATTCGACGAGGAGAAATTAACAAATCATAAAGTATATGTTGTCAGTG 317

Db 264 ATGATGATGAGGATCTCGTGGCGGTTTGGCACTGCCATGTAAGTTTGGCGAAGC 323

QY 318 GACTGCGATATTAGCAGGTGATGCTTTATTAAGCAATTTGAATTTTCAAGTGA 377

Db 324 AAACGCGATTCTCGTGGCGACGCTTTACAAACGCTGGCGTCTTCGATTTTAAGCGATGC 383

QY 378 TCATA-----GATTAACCTGATGAAGTAAATAAAGTCTACAAACGCTGTCAATAGC 431

Db 384 CGATATGCGGGAAGTGTGCGACCGACAGAAATTTTCGATGATTTCTGAACCTGCGAGCGC 443

QY 432 AAGTGGTCAATGTTGAATGCTGGCGGTCAAAATGTTTAGATATGCAAAAGCGAAGGCCAACCC 491

Db 444 CASTGGTATGCGGGAATGTCGGTGGTCAAGCANTTAGATTTAGACGCGGGAAGGCNAACA 503

